

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: April 19, 2002, 21:44:51 ; Search time 76.69 Seconds

(without alignments)  
339.024 Million cell updates/sec

Title: US-09-840-787-19

Perfect score: 1828

Sequence: 1 MADQDAGISPLQMVASGT.....TYEFGKSFQRLNODRLGCG 351

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
A.Geneseq\_1101.\*  
1: /SID2/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1828	100.0	351	21	Human uncoupling P
2	1801	98.5	365	21	Human ORFX ORF2547
3	423.5	23.2	378	21	Arabidopsis thalia
4	423.5	23.2	378	21	Arabidopsis thalia
5	423.5	23.2	413	21	Arabidopsis thalia
6	423.5	23.2	413	21	Arabidopsis thalia
7	374	20.5	318	21	Arabidopsis thalia
8	374	20.5	318	21	Arabidopsis thalia
9	343	18.8	331	21	Arabidopsis thalia
10	342	18.7	331	21	Arabidopsis thalia
11	335	18.3	290	21	Arabidopsis thalia

12	334	18.3	290	21	AA622078	Arabidopsis thalia
13	322	17.7	278	21	AA633094	Arabidopsis thalia
14	322	17.6	278	21	AA622079	Arabidopsis thalia
15	320	17.5	289	18	AA617054	Blackcurrant RIB7
16	311	17.0	335	21	AA606090	Arabidopsis thalia
17	299	16.4	305	21	AA606091	Arabidopsis thalia
18	293	16.0	318	22	AA641505	Human polypeptide
19	286.5	15.7	300	21	AA613086	Arabidopsis thalia
20	286.5	15.7	300	21	AA637830	Arabidopsis thalia
21	286.5	15.7	300	21	AA639526	Arabidopsis thalia
22	285	15.6	330	21	AA628434	Arabidopsis thalia
23	285	15.6	332	21	AA628433	Arabidopsis thalia
24	284.5	15.6	366	22	AA600938	Human bone marrow
25	284	15.5	459	21	AA628051	Arabidopsis thalia
26	284	15.5	502	21	AA628050	Arabidopsis thalia
27	284	15.5	617	21	AA628049	Arabidopsis thalia
28	280.5	15.3	266	21	AA628435	Arabidopsis thalia
29	280	15.3	289	21	AA650383	Human uncoupling P
30	280	15.3	289	22	AA660113	Human transport pr
31	271	14.8	305	22	AA640265	Human adenine nucl
32	271	14.8	447	20	AA605713	Human adenine nucl
33	271	14.8	674	20	AA605712	Human ORFX ORF3011
34	271	14.8	708	21	AA643247	Human uncoupling P
35	268.5	14.7	304	21	AA650384	Human polypeptide
36	269.5	14.7	377	22	AA642051	Human polypeptide
37	258.5	14.1	352	21	AA630070	Arabidopsis thalia
38	257	14.1	272	21	AA642980	Human ORFX ORF2744
39	254	13.9	311	21	AA650382	Human uncoupling P
40	254	13.9	311	22	AA639173	Human polypeptide
41	251	13.7	312	21	AA623665	Arabidopsis thalia
42	250.5	13.7	323	21	AA650419	Human uncoupling P
43	250	13.7	291	21	AA630071	Arabidopsis thalia
44	247.5	13.5	268	22	AA639719	Human polypeptide
45	246.5	13.5	310	22	AA693001	Human protein sequ

## ALIGNMENTS

RESULT 1	AA650378	standard: Protein: 351 AA.
ID	AA650378	
AC	AA650378	
XX		
DE	12-MAR-2001 (first entry)	
XX		
XX		Human uncoupling protein #1.
KW	Human; uncoupling protein; immunosuppressive; antiarthritic;	
KW	antirheumatic; antiproliferative; cardiant; vasotrophic;	
KW	cerebroprotective; neuroprotective; antibacterial;	
KW	gastrointestinal; nephrotropic; gynaecological; vulnery;	
KW	gene therapy; cancer; wound; infectious disease; thrombosis;	
KW	infertility.	
OS	Homo sapiens.	
PN	WO200061614-A2.	
PD	19-OCT-2000.	
XX		
XX		
PF	06-APR-2000; 2000WO-US09534.	
XX		
PR	09-APR-1999; 99US-0128701.	
PR	08-JUL-1999; 99US-0142821.	
PR	18-AUG-1999; 99US-0149448.	
PR	12-NOV-1999; 99US-0164751.	
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Ruben SM, NI J, Komatsoulis G, Rosen CA, Soppet DR;	
XX		

DR WPI; 2000-656322/63.  
 DR N-PSDB; AAC90452.  
 XX  
 PT Uncoupling proteins and nucleic acid sequences encoding them, useful  
 PT for detecting, preventing and treating proliferative, neurological,  
 PT immune system, cardiovascular and gastrointestinal disorders -  
 XX  
 PS Claim 11; Page 316-317; 343pp; English.  
 XX  
 CC The present sequence is a human uncoupling protein. The nucleotide  
 CC sequences encoding the uncoupling proteins may be used for  
 CC the detection of various disorders such as cancer, for chromosome  
 CC identification, as chromosome markers and for numerous other diagnostic  
 CC or research purposes. The uncoupling protein encoded by the nucleotide  
 CC sequences may be used to treat disorders such as neural, immune,  
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
 CC renal and proliferative disorders, wounds, infectious diseases,  
 CC thrombosis, arthritis, and infertility.  
 CC  
 XX  
 SQ Sequence 351 AA;

Query Match 100.0%; Score 1828; DB 21; Length 351;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-186;  
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MADQDPAGISPLQOMWASGTGAVVTSLEMTPLDVVKVRLQSORPSMASLMPSSRLMSTLS 60  
 DB 1 madqdpagisplqgmwasgtgavvtslmtpldvkvrlqsgfpmaselmssrlmstls 60  
 OY \* 61 YTKMKCLLYNGVLEPIYLCPNGARCATWQDPRTFTGTMDAVKIVRHGRTIWSGLP 120  
 DB 61 ytkmckllycngvlepylcpngarcattwqdprrtftgtmdatvkivrhgtrtlwsqglp 120  
 OY 121 ATLMTVPATAIETFAVYDQKAFKALGSRALTSIDYAPVAVAGALARLGTWTVISPELEMRK 180  
 DB 121 atlmtvpataiyftaydqkafkalgrraltsdyapmvagalarlgtvtvisplemrk 180  
 OY 181 LQAHVSYRELGACVRRPAAVAGGWSRLWLGSPALRDVPFSAIYWFNELYKSWINGLR 240  
 DB 181 lqahvsyrelgacvrrpaaavaggsrslwlgwspalrdrvfsaiylwfnelykswinglr 240  
 OY 241 PKQGTSGMSEFVAGISGTYAAVLTLPDVVKTOROVAICAMRAVYNPLHVSSTWLLLR 300  
 DB 241 pkqgtsgmstfvaagisgtyaaavltpdvvktrovaicamraavynplhvsstwlillr 300  
 OY 301 RIRAESGTGKLFAGFLPRITIKAPSCAIMISTYEFKSEFORLNDRLGG 351  
 DB 301 riraesgtkylfagflprlikapscaimistyeffkseffqrlndrllgg 351

RESULT 2  
 AAB42783  
 ID AAB42783 standard; Protein; 365 AA.  
 XX  
 AC AAB42783;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF2547 polypeptide sequence SEQ ID NO:5094.  
 XX

Human, open reading frame; ORFX; detection; cytosolic; hepatotropic;  
 vulnerability; antiparasitic; antiparkinsonian; neurotrophic; neuroprotective;  
 anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiant;  
 immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 hypotensive; dermatological; immunosuppressive; antinflammatory;  
 antiviral; antibacterial; antifungal; antineumatic; antihypertensive;  
 antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 cholesterol ester storage; systemic lupus erythematosus; infection;  
 severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;

KM bone damage; cartilage damage; antinflammatory disease; coagulation;  
 KM thrombosis; contraceptive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2005058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000WO-US08621.  
 XX  
 PR 31-MAR-1999; 9905-0127607.  
 PR 02-APR-1999; 9905-0127636.  
 PR 05-APR-1999; 9905-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shinkets RA, Leach M;  
 XX  
 DR WPI; 2000-602362/57.  
 DR N-PSDB; AAC76992.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 XX  
 PS Claim 11; Page 4282-4283; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytosolic; hepatotropic; vulnerability;  
 CC antiparasitic; antiparkinsonian; neurotrophic; neuroprotective;  
 CC osteopathic; anticonvulsant; antiarthritis; immunosuppressant;  
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
 CC antinflammatory; antibacterial; antiviral; antifungal; antineumatic;  
 CC antihypertensive; antianemic. The sequences can be used for determining  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORFX-associated disorder. The  
 CC nucleic acids can be used to express ORFX proteins in gene therapy  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, immunodeficiency (SCID), AIDS, viral,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance  
 CC coagulation; to inhibit thrombosis; and as a contraceptive.  
 CC  
 XX  
 SQ Sequence 365 AA;

Query Match 98.5%; Score 1801; DB 21; Length 365;  
 Best Local Similarity 96.2%; Pred. No. 1e-183;  
 Matches 351; Conservative 0; Mismatches 0; Indels 14; Gaps 2;

OY 1 MADQDPAGISPLQOMWASGTGAVVTSLEMTPLDVVKVRLQSORPSMASLMPSSRLMSTLS 57  
 DB 1 madqdpagisplqgmwasgtgavvtslmtpldvkvrlqsgfpmaselmssrlmstls 60  
 OY 58 -----SLSTYKKKCLLYCNGVLEPIYLCPNGARCATWQDPRTFTGTMDAVKIVRHGRT 112  
 DB 61 ytklpelstykckllycngvlepylcpngarcattwqdprrtftgtmdatvkivrhgtr 120  
 OY 113 RTLWSGLPALTMTVPATAIETFAVYDQKAFKALGSRALTSIDYAPVAVAGALARLGTWTVIS 172  
 DB 121 rtlwsqglpaltmtvpataiyftaydqkafkalgrraltsdyapmvagalarlgtvtvis 180  
 OY 173 PLELMRTKIQAHVSYRELGACVRRPAAVAGGWSRLWLGSPALRDVPFSAIYWFNELYKSW 226  
 DB 181 plelmrtkigahvsyrelgacvrrpaaavaggsrslwlgwspalrdrvfsaiylwfnelyksw 240

OY 227 FNTFLKSWMLNGLRPKDQTSVGSMEFVAGGISTVAAVLTPEDVYKTKQROVALGAMEAVR 286  
|||||  
Db 241 fnyelkxwnglrirpdkcsvgsfvaegglsgltvaavlllpdvvktqfvalgameavr 300  
OY 287 VNPLHWDSTWLLLRIRRAESGKGLFAGFLPRIRKAAPSCAIMISYERKSFQRLNOD 346  
|||||  
Db 301 vnplhwdstwlllrirraesgkglfagflprirkaapscaimisyltqkslfrlqnd 360  
OY 347 RLGG 351  
|||||  
Db 361 rlllg 365  
RESULT 3  
AAG21416  
ID AAG21416 standard; Protein; 378 AA.  
AC AAG21416;  
XX 17-OCT-2000 (first entry)  
DT  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23965.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
PD  
XX 06-SEP-2000.  
PF  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139889.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140693.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142134.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 22-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147182.  
PR 05-AUG-1999; 99US-0147260.

Query Match	23.28;	Score 423.5;	DB 21;	Length 378;
Best Local Similarity	32.08;	Pred. NO. 1.4e-36;		

Matches	117;	Conservative	57;	Mismatches	137;	Indels	55;	Gaps	9;
QY	13	QOMVASGTAVVTSLFMTPLDVYKVRISORPSMASELMPPSRILW-----SLSTYTKWK	65						
Db	28	erfsaagaavlsavlndpldvkkrllgdaagmsyshpmslsgmatfpgpnmfadlr	87						
QY	66	CLLYC--NGVLEPLYLICPNARCATFQDPTRFTGTMDAFVKIVRHEGTRTLMSGIPATL	123						
Db	88	cspscaraagvegtsicp-----pdctgkgtfdvfkllirgeqigrirwrgtnagll	138						
QY	124	VMTVPATATVFTAYD-----QLKAFILGRALTSDLVAPWAGLARIGRTVTSPELMRT	179						
Db	139	alevmwgiylpdydmfnrlnleelstrekepanntfcyptvgaaglaslactvypidlatr	198						
QY	180	KLDAQHVSRELGAC-----VRTA---VAOGMRSIMLWGMPALRDV	219						
Db	199	rmga-----fkeakagykppgvfkllvgvfsevrtnanlesslhnryglwrglgaqlardv	254						
QY	220	PFSALYFWVEVLKWSLNLGRPKDQTSVGM---SPVAGISGTVAAVLTLPEDVYKTRQ	276						
Db	255	pfaalcwstlepkikrllgvagndcnlyvgfatisagfiagslaaactcpldvartlrq	314						
QY	277	VALGMAEAVRNPDLHVDSTWLLRRIRAESGTGGLFAGFLPRRIKAAPSCAIMISTYFEG	336						
Db	315	lekdpjralmm-----tlrrqlilewrdgmgriifmgmgprvaragpsvgilvsvfyev	368						
QY	337	KSEFOR	342						
Db	369	kyvlhr	374						
RESULT	4								
AAG40152									
ID	AAG40152	standard; Protein: 378 AA.							
XX									
AC	AAG40152;								
XX									
DT	18-OCT-2000	(first entry)							
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 49782.							
XX									
KW	Protein identification: signal transduction pathway; metabolic pathway;								
KW	hybridisation assay; genetic mapping; gene expression control; promoter;								
XX	termination sequence.								
XX									
OS	Arabidopsis thaliana.								
XX									
FN	EP1033405-A2.								
XX									
PD	06-SEP-2000.								
XX									
FE	25-FEB-2000;	2000EP-0301439.							
XX									
PR	25-FEB-1999;	99US-0121825.							
PR	05-MAR-1999;	99US-0123180.							
PR	09-MAR-1999;	99US-0123548.							
PR	23-MAR-1999;	99US-0125788.							
PR	25-MAR-1999;	99US-0126264.							
PR	29-MAR-1999;	99US-0126785.							
PR	01-APR-1999;	99US-0127432.							
PR	06-APR-1999;	99US-0128234.							
PR	08-APR-1999;	99US-0128714.							
PR	16-APR-1999;	99US-0129845.							
PR	19-APR-1999;	99US-0130077.							
PR	21-APR-1999;	99US-0130449.							
PR	23-APR-1999;	99US-0130510.							
PR	23-APR-1999;	99US-0130891.							
PR	28-APR-1999;	99US-0131449.							
PR	30-APR-1999;	99US-0132048.							
PR	04-MAY-1999;	99US-0132407.							
PR	05-MAY-1999;	99US-0132484.							
PR	05-MAY-1999;	99US-0132485.							
PR	06-MAY-1999;	99US-0132486.							

PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140822.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145219.  
PR 26-JUL-1999; 99US-0145224.  
PR 27-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147302.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148177.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149422.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.

PR	23-MAR-1999;	9905-01253788;
PR	25-MAR-1999;	9905-01262664;
PR	29-MAR-1999;	9905-01262785;
PR	01-APR-1999;	9905-01274652;
PR	06-APR-1999;	9905-01282634;
PR	08-APR-1999;	9905-01282845;
PR	16-APR-1999;	9905-01287814;
PR	19-APR-1999;	9905-01300777;
PR	21-APR-1999;	9905-01303449;
PR	23-APR-1999;	9905-01305010;
PR	23-APR-1999;	9905-01308911;
PR	28-APR-1999;	9905-01314449;
PR	30-APR-1999;	9905-01323048;
PR	30-APR-1999;	9905-01324077;
PR	04-MAY-1999;	9905-01324484;
PR	05-MAY-1999;	9905-01324855;
PR	06-MAY-1999;	9905-01324867;
PR	06-MAY-1999;	9905-01324876;
PR	07-MAY-1999;	9905-01328663;
PR	11-MAY-1999;	9905-01342518;
PR	14-MAY-1999;	9905-01343246;
PR	14-MAY-1999;	9905-01343219;
PR	14-MAY-1999;	9905-01343221;
PR	14-MAY-1999;	9905-01343760;
PR	18-MAY-1999;	9905-01343768;
PR	19-MAY-1999;	9905-01343941;
PR	20-MAY-1999;	9905-01353124;
PR	21-MAY-1999;	9905-01353523;
PR	24-MAY-1999;	9905-01356021;
PR	25-MAY-1999;	9905-01356021;
PR	27-MAY-1999;	9905-01363592;
PR	28-MAY-1999;	9905-01363782;
PR	01-JUN-1999;	9905-01373222;
PR	03-JUN-1999;	9905-01375028;
PR	04-JUN-1999;	9905-01375052;
PR	08-JUN-1999;	9905-01377624;
PR	08-JUN-1999;	9905-01383604;
PR	10-JUN-1999;	9905-01383847;
PR	14-JUN-1999;	9905-01393119;
PR	16-JUN-1999;	9905-01394532;
PR	16-JUN-1999;	9905-01394543;
PR	17-JUN-1999;	9905-01394924;
PR	18-JUN-1999;	9905-01394942;
PR	18-JUN-1999;	9905-01394651;
PR	18-JUN-1999;	9905-01394652;
PR	18-JUN-1999;	9905-01394653;
PR	18-JUN-1999;	9905-01394577;
PR	18-JUN-1999;	9905-01397550;
PR	18-JUN-1999;	9905-01397653;
PR	21-JUN-1999;	9905-01398177;
PR	22-JUN-1999;	9905-01398939;
PR	22-JUN-1999;	9905-01403553;
PR	23-JUN-1999;	9905-01403554;
PR	24-JUN-1999;	9905-01406955;
PR	28-JUN-1999;	9905-01408223;
PR	29-JUN-1999;	9905-01409591;
PR	30-JUN-1999;	9905-01412887;
PR	01-JUL-1999;	9905-01418482;
PR	01-JUL-1999;	9905-01421544;
PR	02-JUL-1999;	9905-01420555;
PR	06-JUL-1999;	9905-01429290;
PR	09-JUL-1999;	9905-01428203;
PR	12-JUL-1999;	9905-01429270;
PR	13-JUL-1999;	9905-01435422;
PR	14-JUL-1999;	9905-01436224;
PR	15-JUL-1999;	9905-01444005;

PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-015823.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

DE Arabidopsis thaliana protein fragment SEQ ID NO: 49781.  
XX Protein identification: signal transduction pathway; metabolic pathway;  
KM hydrolisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
OS EPI033405-A2.  
PN  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143544.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144003.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.

```
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154739.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 23.2%; Score 423.5; DB 21; Length 413;

Best Local Similarity 32.0%; Pred. No. 1.6e-36;

Matches 117; Conservative 57; Mismatches 137; Indels 55; Gaps 9;

```
OY 13 QOMVASTGAVTSLFMTPLDVKVRLQSORPSMASLMPSSRRM-----SLSTYKWK 65
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 ervfsaagaavlsavtlnpdlvkvtrlqagaagsysphlsnslgmaffgpmmmfaldl 122
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 66 CLVVC-NGVLEPLYLCPNGARCTWFODPTRFGTMDAFYKLYRHGCTRTMGLPATL 123
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 123 cspscaraevgevtslcp-----pdcfgykgtldvfkllirgsglrlwrgtaql 173
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 124 VMTVPATAIYFTAYD---OLKAFLCGRALTSIDLYAPMVAGALRLCTVVISPLELMRT 179
      : : | | | | : : : : : : : : : : : : : : : : : : : : : : : : :
Db 174 alavpmwgilylpfydmfnrlleelsrckapamfcvptvagslarslactvcypidart 233
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 180 KLGQAHYSYRELAC-----VRTA-----VAQCGWRSLMLGMPFALRDV 219
      : : | | : : : : | | | | : : : : : : : : : : : : : : : : : : :
Db 234 rmga-----fkeakagvvpqpyfkllygvfsevrtaannlesslhnyrglwrqlgaqlardv 289
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 220 PFSALYWFNYELVKSWMGLRPRKDOTSVGM---SFVAGGISGTVAAVLTLPFDVVKTORQ 276
      | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 290 pfsalcwstleplkrllygvagndtnlygvfgatfsagflagslaaatacpldvarcrrg 349
```

```
OY 277 VALGAEAVRNPLHVDSTWLLLRIRAESGTGCLPAGFLPRITKAAPSALMISTYERG 336
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 350 iekdpgrlmm-----ttrgcllewrtdgmgrlfmgmgpvaragsvgtvsvfgev 403
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 337 KSFQQR 342
      | | | |
Db 404 kylvhr 409
      : : : :

RESULT 7
AAG21417
ID AAG21417 standard; Protein; 318 AA.
XX
AC AAG21417;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23966.
XX
KW 'Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126264.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 06-APR-1999; 99US-0128234.
XX
PR 08-APR-1999; 99US-0128714.
XX
PR 16-APR-1999; 99US-0129845.
XX
PR 19-APR-1999; 99US-0130077.
XX
PR 21-APR-1999; 99US-0130449.
XX
PR 23-APR-1999; 99US-0130510.
XX
PR 28-APR-1999; 99US-0130891.
XX
PR 30-APR-1999; 99US-0131449.
XX
PR 30-APR-1999; 99US-0132048.
XX
PR 04-MAY-1999; 99US-0132407.
XX
PR 04-MAY-1999; 99US-0132484.
XX
PR 05-MAY-1999; 99US-0132485.
XX
PR 06-MAY-1999; 99US-0132486.
XX
PR 07-MAY-1999; 99US-0132487.
XX
PR 11-MAY-1999; 99US-0132863.
XX
PR 14-MAY-1999; 99US-0134218.
XX
PR 14-MAY-1999; 99US-0134219.
XX
PR 14-MAY-1999; 99US-0134221.
XX
PR 14-MAY-1999; 99US-0134370.
XX
PR 18-MAY-1999; 99US-0134768.
XX
PR 19-MAY-1999; 99US-0134941.
XX
PR 20-MAY-1999; 99US-0135124.
XX
PR 21-MAY-1999; 99US-0135353.
XX
PR 24-MAY-1999; 99US-0135629.
XX
PR 25-MAY-1999; 99US-0136021.
XX
PR 27-MAY-1999; 99US-0136392.
XX
PR 28-MAY-1999; 99US-0136782.
XX
PR 01-JUN-1999; 99US-0137222.
XX
PR 03-JUN-1999; 99US-0137528.
XX
PR 04-JUN-1999; 99US-0137502.
XX
PR 07-JUN-1999; 99US-0137724.
XX
PR 08-JUN-1999; 99US-0138094.
XX
PR 10-JUN-1999; 99US-0138540.
XX
PR 10-JUN-1999; 99US-0138847.
```

PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140355.  
PR 28-JUN-1999; 99US-0140695.  
PR 29-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140891.  
PR 01-JUL-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142155.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142890.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149829.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159684.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161992.  
PR 29-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

## Query Match

20.5%; Score 374; DB 21; Length 318;  
Best Local Similarity 32.8%; Pred. No. 2.2e-31;

Matches 103; Conservative 43; Mismatches 120; Indels 48; Gaps 8;

QY 58 SLSYTKWKCLAYC--NGVLEPLYLCPNGARCAATWQDPTRFRTGMDAVKIVRHGCTRTL 115  
DB 20 mmfifaditrcpscacaragvegtvscp-----pdcfgfykgtfdvftkiltirgeglgrl 70



```

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 23-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

```

```

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 29-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.5%; Score 374; DB 21; Length 318;
Best Local Similarity 32.8%; Pred. No. 2, 2e-31;
Matches 103; Conservative 43; Mismatches 120; Indels 48; Gaps 8;

OY 58 SLSTYKWKCLLYC--NGVLEPLYLCPNGARCATWFOPTFRFTGMDAFYKIVRHGSTRLL 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 20 nmffadrlrcpsccaragvegtvsicp-----pdcftgyxftdvfkilrtgeqlgrl 70

OY 116 WSGLPATLWTPVPAATAYTAYD---OLKAPLCGRALSLDLYAPVAGALRLGTVPI 171
   | | | | : : | | | | | : : : : : : : : : : : : : : : : : : : : :
DB 71 wrtgnglglavpwngvlylptfdmfrlleejsrekapamtfcvplvagslarslactvc 130

OY 172 SPLELMRTKLQAOHVSRYRELGAC-----VPTA---VAQGGMRSLMLGM 211
   | : | | : : | | : : | : : : : : : : : : : : : : : : : : : : : :
DB 131 ypldlatrtmga-----fkakagvdkppgyfktlvgyfsevrtaannlesslmyrqlwrgl 186

OY 212 GPTALBDVPFSALWPNELVKSWLNGLRPKDQTSVM--SPVAGGISGTVAAVLTLPF 268
   | : | | | | | : : : : : : : : : : : : : : : : : : : : : : : :
DB 187 gqglardvpfsalcwstlepkikrllvgandtnlvgvfagfsagflagslaaatcpl 246

OY 269 DVVKTORQVALGAMAVRNPLHVDSTWLLLRIRAESGTRKGLFAGFLPRIIKAPSCAI 328
   | : | : | : | : : : : : : : : : : : : : : : : : : : : : : : :
DB 247 dvartrqlekqdgrralnm-----ltrqtlilevrldgmrglimgmprvaragsvyl 300

OY 329 MISTYEFGRKSPFOR 342
   : : | | | | | : : : : : : : : : : : : : : : : : : : : :
DB 301 vvsfyevkkyvlhr 314

RESULT 9
AAG43092
ID AAG43092 standard; Protein; 331 AA.
XX
AC AAG43092;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53820.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
FD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 29-MAR-1999; 99US-0126264.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.

```

PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141847.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144332.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147933.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149829.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151067.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.

```
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.
```

Query Match 18.8%; Score 343; DB 21; Length 331;  
Best Local Similarity 27.8%; Pred. No. 4.7e-28;  
Matches 99; Conservative 58; Mismatches 133; Indels 66; Gaps 7;

```
OY 3 DDPAGISPLDQ-----MVASGTGAVTSLFMTPLDVKVRLOQSRPMSASELMPSS 54.  
DB 22 dhpaiivpaquttlkfmglmwagslsgsvelhnamfpvdtvkthmgalt----- 70  
OY 55 RLMSLSYRKWKCLYCNGLVEPLVLCPPNARCATWFQODPTREFTGMDARVKIVRHGRT 114  
DB 71 -----sc-----plkpiqirgastrsliktdgsaa 94  
OY 115 LMSGLPATLVMTVPATAIYFAVDLQKAFGLGRLTSDIYAPMVAGALRLGTVTYISPL 174  
DB 95 lyrgtwanglqagpahavyfsfyevskkfllsg-gpnmsaahaisgvtffissdavidfpm 153  
OY 175 ELMPKRLQAHYRSLGACVNTAVAGGWSLWLGMEFTALRDVPFSALYFNTELYKS 234  
DB 154 dmvkqrlqigntkygvwdcklrvreegfgafyasytvlvmapflavhlttyeaavr 213  
OY 235 WLNGLRP-----KDQTSVGSMSFVAGISGTVAVALTLPEFDVVKTOROVALGAMEAVERN 288  
DB 214 glremhpehayaadeegwilyatagaagajlaaavtupldvckcqldq-cgvgcgcdrtk 272  
OY 289 PLHVDSTWLLMRIRAESGTGKGLFAGFLPRIRIKAPSCAIMISTVEFGKSPFORLN 344  
DB 273 sssisd---vftlivkkdgyrglrgawlprmlfhapaaicwstvetvksffgdln 325
```

## RESULT 10

AAG22077 AAG22077 standard; Protein; 331 AA.

AAG22077;

17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 24869.

Protein identification: signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

25-FEB-1999; 99US-0121825.

05-MAR-1999; 99US-0123180.

09-MAR-1999; 99US-0123548.

```
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0125264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134376.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0138459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.
```



termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EPI03405-A2.  
XX XX 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 23-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0128845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138099.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143622.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145112.  
PR 23-JUL-1999; 99US-0145114.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146386.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148365.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149175.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0150566.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.

```
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 18.3%; Score 335; DB 21; Length 290;  
Best Local Similarity 27.7%; Pred. No. 2.8e-27;  
Matches 93; Conservative 58; Mismatches 127; Indels 58; Gaps 6;

```
OY 15 MVASGTGAVVTSLEFMTPLDVVKVRLQSQRPMSASELMPSSRLMSLSYTKMKCLLYCNGVL 74
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 mvagslsgsvelhamfpvdtvktkmgalf-----sc----- 31
OY 75 EPLYLCPNGARCAWTFQDPRTFTGTMDAFVKIVRHEGRTLTLSGCLPATLVMTVPATAIYF 134
    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 32 -----pikpigtqafsrilkttdgpsalryglawmg1gaqpahavyf 73
OY 135 TAYDOLKAFICGRALTSDLVAPVAGALARLIGTVTVISPLELMRTLOAHVSYRELGC 194
    : | : | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 74 slveavkktf1sg-gnpnaaah1sgvfal1ssdavfcfpmawkkq1q1gnqtkykwdc 132
OY 195 VRTAVAGOGWRSLMLGMPALRDVFPFSA1LWMEVELVKSWM1GLRP-----KDQTSVG 248
    : : : : : : : : : | : | : | : | : | : | : | : | : | : | : | : |
Db 133 ikvtteegfgyaayrvtlvmapftavhfttlyeavkriglem1pehavgeedeegwl 192
OY 249 MSFVAGISGTVAVLTLFPDVVKTKTORQVALGMEAVRNP1LHVDSITWLLRIRAESGT 308
    : | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 193 lyatagaaglaaavt1pdrvktq1q-cqyvcgcdrfkss1sd---vftlvkddy 248
OY 309 KGLEAGFLPRITKAAPSCA1MISTYERGKSFQRLN 344
    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 249 rg1argw1pml1thapaal1csw1cy1vksf1gdl 284
```

```
RESULT 12
AAG22078
ID AAG22078 standard; Protein; 290 AA.
XX
AC AAG22078;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24870.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
```

```

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145199.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

```

```

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155658.
PR 26-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158263.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161997.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 18.3%; Score 334; DB 21; Length 290;

Best Local Similarity 27.7%; Pred. No. 3, 6e-27; Mismatches 127; Indels 58; Gaps 6;

```

Matches 93; Conservative 58; Mismatches 127; Indels 58; Gaps 6;

QY 15 MWASGTGAVVTSLEFMPBLDVVYKVLQSRPSMASLMPSSRLMSLSTKWKCLLYNGVYL 74
Db 1 mwagslagsvvehnamfvdvclthmgqlr-----sc----- 31

QY 75 EPLYLCPNGARCATWFPDPRFRFTGTMDFYKIVRHGECTRTLSGLPATLVMTVPATAIYF 134
Db 32 -----pikpigitqratsilktqpsalsygiwamlgagpahayf 73

QY 135 TAYDOLKAFPLGRALTSGLAPVWAGALARLGTIVTISPLELRTKLOAOHVSRELIGAC 194
Db 74 sfyevskkfllsg-gnpnnsahalsgyfatissdavftfpmamvkgqligngtgykgvdc 132

```



PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0155658.
PR	29-SEP-1999;	99US-0155656.
PR	04-OCT-1999;	99US-0155717.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160987.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	28-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match		
Best Local Similarity	17.7%;	Score 323; DB 21;
	27.6%;	Pred. No. 5,1e-26;
		Length 278;

Matches	89; Conservative	57; Mismatches	118; Indels	58; Gaps
QY	29	MTPLDVKVKELQSRPMSASELMPSSRLMSLSTYTKWKCLLYCNGVLEPLYPENGCARCAT	88	
Db	3	mfpvdvckthmgqlr-----sc-----	19	
QY	89	WFDPPRFETCTMAFAFKIVRHEGSTRILMSGPATVIVYVATATITVTAADLKAFICGKA	148	
Db	20	----pllpilgylrqafslskldqpsalylgicwamgilaqapahavtyslyevskkfils-g	74	
QY	149	LTSDELVAWMAGALARLGVTVISPELMLKTKLOAOHVSYSRELACVRAVAQGWRSIM	208	
Db	75	npmsaahaisgyfatissdavftpmduvkkqrlqigngykygwdcikrvteeegfgatcy	134	
QY	209	LGNGPFRALNRVPPSALYWFMYELVKSWLNGLRP-----KDQTSVGMSTVAGISCTVAA	262	
Db	135	asyltvlmaphlavhftlyeavkrglremlpbhavgaadeegwlliyataaaaglaa	194	
QY	263	VLLPFDVYKTRQVVALGMAEAVRVNPLHVDSTWLLRLRIRAESGTGKGFAGFLPRIKA	322	
Db	195	avtlpdlvkvrtq-qcgyvcgdcdrfkssissd---vfrtlvkkdygrlgarwlpmlfh	250	
QY	323	APSCALMISTYEFGRKSFQRLN	344	
Db	251	apaaalcwstlyeclvksffiqdln	272	
RESULT 14				
ID	AAG22079 standard; Protein; 278 AA.			
XX	AAG22079;			
XX	17-OCT-2000 (first entry)			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 24871.			
XX				
KW	Protein identification; signal transduction pathway; metabolic pathway;			
KW	hybridisation assay; genetic mapping; gene expression control; promoter;			
XX	termination sequence.			
XX	Arabidopsis thaliana.			
PN	EPI033405-A2.			
XX				
PD	06-SEP-2000.			
XX				
PF	25-FEB-2000; 2000EP-0301439.			
XX				
PR	25-FEB-1999;	99US-0121825.		
PR	05-MAR-1999;	99US-0123180.		
PR	09-MAR-1999;	99US-0123548.		
PR	23-MAR-1999;	99US-0125788.		
PR	25-MAR-1999;	99US-0126264.		
PR	29-MAR-1999;	99US-0126785.		
PR	01-APR-1999;	99US-0127462.		
PR	06-APR-1999;	99US-0128234.		
PR	08-APR-1999;	99US-0128714.		
PR	16-APR-1999;	99US-0129845.		
PR	19-APR-1999;	99US-0130077.		
PR	21-APR-1999;	99US-0130449.		
PR	23-APR-1999;	99US-0130510.		
PR	23-APR-1999;	99US-0130891.		
PR	28-APR-1999;	99US-0131449.		
PR	30-APR-1999;	99US-0132048.		
PR	30-APR-1999;	99US-0132407.		
PR	04-MAY-1999;	99US-0132484.		
PR	05-MAY-1999;	99US-0132485.		
PR	06-MAY-1999;	99US-0132486.		
PR	07-MAY-1999;	99US-0132487.		
PR	11-MAY-1999;	99US-0132863.		
PR	14-MAY-1999;	99US-0134256.		
PR		99US-0134218.		

PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134321.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134766.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140699.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142153.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145109.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156556.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 22-OCT-1999; 99US-0160999.



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 19, 2002, 21:47:27 ; Search time 69.69 Seconds  
(without alignments)  
383.660 Million cell updates/sec

Title: US-09-840-787-19

Perfect score: 1828  
Sequence: 1 MADQDPAGISPLQQWVASGT.....TYEFGKSFFORLNDRLG 351

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: PIR\_60:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1828	100.0	351	2 T43493	hypothetical prote
2	648.5	35.5	328	2 T19322	hypothetical prote
3	486.5	26.6	335	2 T50393	probable mitochon
4	474.5	26.0	366	2 S64589	hypothetical prote
5	348	19.0	358	2 D84901	hypothetical prote
6	343	18.8	331	2 T00582	probable mitochon
7	312	17.1	781	2 A86305	hypothetical prote
8	300.5	16.4	378	2 T02895	hypothetical prote
9	286	15.6	332	2 T47703	Ca-dependent solut
10	284.5	15.6	902	1 S54495	probable carrier p
11	281.5	15.4	310	2 S69050	probable membrane
12	277.5	15.2	348	2 D84798	probable mitochon
13	273.5	15.0	358	2 T45934	hypothetical prote
14	272	14.9	312	2 T25089	hypothetical prote
15	268.5	14.7	368	2 S54524	hypothetical prote
16	264	14.4	702	2 T16533	hypothetical prote
17	260.5	14.3	302	2 T38879	probable mitochon
18	260.5	14.3	338	2 T40968	probable phosphate
19	259.5	14.2	307	2 S60949	hypothetical prote
20	255.5	14.2	309	2 T29225	mitochondrial solu
21	258.5	14.1	352	2 T01729	mitochondrial solu
22	257	14.1	475	2 T50686	peroxisomal Ca-dep
23	256	14.0	303	2 T39149	probable RNA splic
24	254	13.9	349	2 B96753	hypothetical prote
25	250.5	13.7	309	2 T48156	hypothetical prote
26	249.5	13.6	325	2 T04273	hypothetical prote
27	249.5	13.6	330	2 T29640	mitochondrial carr
28	249	13.6	304	2 S13533	mRNA splice defec
29	245.5	13.4	377	2 S36081	probable carrier p

30	243.5	13.3	322	2 S57116	probable carrier p
31	237	13.0	347	2 T50990	hypothetical prote
32	235	12.9	343	2 T15253	hypothetical prote
33	233	12.7	330	2 S26596	Graves disease mit
34	233	12.7	373	2 S48451	probable membrane
35	232	12.7	314	2 S51179	mRNA splice defec
36	229	12.5	327	2 S61660	probable mitochon
37	228.5	12.5	447	2 T00435	probable membrane
38	227.5	12.4	339	2 A41677	ADP carrier pr
39	226.5	12.4	284	2 S45458	uncoupling protein
40	226.5	12.4	313	2 T05577	probable mitochon
41	223	12.2	345	2 T37576	hypothetical prote
42	223	12.2	415	2 T48171	Btl protein precu
43	221	12.1	436	2 J01459	probable membrane
44	220	12.0	326	2 S57544	hypothetical prote
45	219.5	12.0	357	2 S46793	hypothetical prote

## ALIGNMENTS

## RESULT 1

T43493  
hypothetical protein DKFZp434C119.1 - human

C:Species: Homo sapiens (man)

C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000

C:Accession: T43493

R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, December 1999

A:Reference number: 22516

A:Accession: T43493

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-351 <AAA>

A:Cross-references: EMBL:AL133584

A:Experimental source: adult testis; clone DKFZp434C119

C:Genetics:

A>Note: DKFZp434C119.1

Query Match	100.0%	Score 1828	DB 2	Length 351
Best Local Similarity	100.0%	Pred. No. 3	1e-149	
Matches 351	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MADQDPAGISPLQQWVASGTGAVVTSLEFMPLDVVKVRLQSQSPMSASLMPSSRLMSLS	60	
DB	1	MADQDPAGISPLQQWVASGTGAVVTSLEFMPLDVVKVRLQSQSPMSASLMPSSRLMSLS	60	
QY	61	YTKMKCLYCNGLVLEPLICPGNGARCATWPDPTREFTGTMDAFYKIVRHGSTRILMSGLP	120	
DB	61	YTKMKCLYCNGLVLEPLICPGNGARCATWPDPTREFTGTMDAFYKIVRHGSTRILMSGLP	120	
QY	121	ATLVMTVPATAYFTAYDOLKAFICGRALTSGLYAPMVAGALRLGTVTVISPELMRTK	180	
DB	121	ATLVMTVPATAYFTAYDOLKAFICGRALTSGLYAPMVAGALRLGTVTVISPELMRTK	180	
QY	181	LOAQHVSTRELACVTRTAAGGMSLWLNCPPTLRVPVSALYWFVELYKSWLNGLR	240	
DB	181	LOAQHVSTRELACVTRTAAGGMSLWLNCPPTLRVPVSALYWFVELYKSWLNGLR	240	
QY	241	PKQOTSVGMSFVAGISGTVAAVLTLPEPDVYKTORQVALGMAEAVRNPVLHVDSTWLLR	300	
DB	241	PKQOTSVGMSFVAGISGTVAAVLTLPEPDVYKTORQVALGMAEAVRNPVLHVDSTWLLR	300	
QY	301	RIRAESGTGLFAGFLPRITIKAPSCAIIWISTYERFGKSFFORLNDRLG 351		
DB	301	RIRAESGTGLFAGFLPRITIKAPSCAIIWISTYERFGKSFFORLNDRLG 351		
RESULT 2				
T19322		hypothetical protein C16C10.1 - Caenorhabditis elegans		
C:Species: Caenorhabditis elegans				

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Mar-2000  
 C:Accession: T19322  
 R:Lloyd, C.  
 submitted to the EMBL Data Library, November 1994  
 A:Reference number: Z19108  
 A:Accession: T19322  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-328 <WIL>  
 A:Cross-references: EMBL:Z46787; PIDN:CA86739.1; GSPDB:GN00021; CESP:C16C10.1  
 A:Experimental source: clone C16C10  
 C:Genetics:  
 A:Gene: CESP:C16C10.1  
 A:Map position: 3  
 A:Introns: 22/1; 125/2; 179/1; 298/1  
 C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 35.5%; Score 648.5; DB 2; Length 328;  
 Best Local Similarity 39.4%; Pred. No. 4.2e-48;  
 Matches 136; Conservative 63; Mismatches 113; Indels 33; Gaps 6;

QY 10 SPLQOMVASTGAVVTSLEMTPLDVVKVRLQSO-RPSMASELMPSSRLMSLSTYKMKCL 68  
 DB 3 SPCESGKTTNCSGAPSSSCVTPIDVVKIRLQOQTRP-----FPKGECCFY 46

QY 69 YCNQVLEPLVLCPNGARCATWFOPTREFTGTMDAFVKIVHEGSTRMLSGPLATVMTVP 128  
 DB 47 YHNLIMHVCSEVERKPCMYORPCGNFRGTADAIYKIHHEGIRLSLMSLSTYKMALE 106

QY 129 ATATYFAVYOLKAF-----CGRALSDIAXP-----WVAGALRLGTVTVISPLEM 177  
 DB 107 ATATYFTTYNLSLWMLKKKCCRAKSPKMTPEWSAAVAIVARTIAVTVVSPTEMI 166

QY 178 RTKLQAOHVSRELGCACVTRTAQAQGWRSMLGCPALDPEFSALYMFNEVLVSMIN 237  
 DB 167 RTKMSKRLTYHEIGHVRSMAFKGISSFYLGWFTMLRDIPFSQITVAGIDLFTNTIQ 226

QY 238 GLRPKOOTSVGMSFVAGGISTYAAVLTLPFDVVKTOROVALGAMAVRNPLHVDSTWL 297  
 DB 227 RRGQPDHNPVSVSGAAGVVASIFTHPFVYKTNCOIRIGG---SIDMKNKST-T 281

QY 298 ILNRIRAESGSKGLFAGFLPRITKAPSCAIMSTYEFKSEFOR 342  
 DB 282 VIKDMYHRSIGISAFSSGIVPRLVKVSFSCAIMSTYEFKFLPÖK 326

RESULT 3  
 T50393  
 probable mitochondrial carrier protein [imported] - fission yeast (Schizosaccharomyces f.  
 C:Species: Schizosaccharomyces pombe  
 C>Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jun-2000  
 C:Accession: T50393  
 R:Beck, A.; Botz, K.; Reinhardt, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, January 1999  
 A:Reference number: Z25067  
 A:Accession: T50393  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-335 <BEC>  
 A:Cross-references: EMBL:AL136535; PIDN:CA86434.1; GSPDB:GN00067; SPDB:SPBP23A10.06  
 A:Experimental source: strain 972h(-); clone pl p23A10  
 C:Genetics:  
 A:Gene: SPDB:SPBP23A10.06  
 A:Map position: 2

Query Match 26.6%; Score 486.5; DB 2; Length 335;  
 Best Local Similarity 36.1%; Pred. No. 3.4e-34;  
 Matches 127; Conservative 48; Mismatches 112; Indels 65; Gaps 9;

QY 9 ISPLQOMVASTGAVVTSLEMTPLDVVKVRLQSO-RPSMASELMPSSRLMSLSTYKMKCL 68  
 DB 107 ATATYFTTYNLSLWMLKKKCCRAKSPKMTPEWSAAVAIVARTIAVTVVSPTEMI 166

DB 22 VKPIAKMLSACVSVITTLTVTPLDVVKTRLOSESISQYSSSTPISAKIL----- 72

QY 69 YCNQVLEPLVLCPNGARCATWFOPTREFTGTMDAFVKIVHEGSTRMLSGPLATVMTVP 128  
 DB 73 ---GKGRPA-----PKPLGQPVSGLYQIARHBSVRLMGLVPSLTMLP 114

QY 129 ATATYFAVYOLKAFGLGRALSDIAXP-----WVAGALRLGTVTVISPLEMRTKLQAOH 185  
 DB 115 ANVVOELGYEOL-----LPLVSDMGEPAAALAGASARTISATIVSPELFRTRVQANG 168

QY 186 VSY-----RELGACY-----RTAAQAGWRSIMLGWCPTRLRQVPSALYMFNEVLKSKL 236  
 DB 169 GHYPPGHAREIANEVVDGLKIMHQGLNLSGVSVTLMRQVPSAFYMSYERIRLFL 228

QY 237 NGLRPKDQT-----SVGMSFVAGGISTYAAVLTLPFDVVKTOROVALGAMAVRN 288  
 DB 229 LG-HPSLQAFSSSQSYKDYINPVSQISGLTATLTQGFQVDSKAKQV----- 276

QY 289 PLHVDSTWLLNRIRAESGSKGLFAGFLPRITKAPSCAIMSTYEFKSEFOR 340  
 DB 277 HGHTLTRGQFMTLMKRGSKALMGKTLPRCVKVAFSCAIMSTYHLYTKYF 328

RESULT 4  
 S64589  
 hypothetical protein YGR257c - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein G9175  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 29-Oct-1999  
 C:Accession: S64589  
 R:Agostoni Carbone, M.L.; Panzeri, L.; Melchiorro, P.; Carignani, G.; Peroli, F.; F.  
 submitted to the Protein Sequence Database, May 1996  
 A:Reference number: S64577  
 A:Accession: S64589  
 A:Molecule type: DNA  
 A:Residues: 1-366 <AO>  
 A:Cross-references: EMBL:Z73042; NID:q1323467; PID:e243687; PID:q1323468; GSPDB:GN000  
 A:Experimental source: strain S288C  
 C:Genetics:  
 A:Gene: MIPS:YGR257c  
 A:Map position: 7R

Query Match 26.0%; Score 474.5; DB 2; Length 366;  
 Best Local Similarity 32.9%; Pred. No. 4.1e-33;  
 Matches 130; Conservative 65; Mismatches 121; Indels 79; Gaps 14;

QY 1 MADDPAGISPL-QOMVASTGAVVTSLEMTPLDVVKVRLQSO-RPSMASELMPSSRLMSL 59  
 DB 1 MSDRNTSLSLTKERMLSAGASVLTSLITLPMQVYRIRLQOQ---QMIPDCS--- 50

QY 60 SYTKWKCLVYCNQVLEPLVLCPNGARCAT-----W-----FQD-----PT 94  
 DB 51 -----CDGAEPVNAVSSGSKKKTFTNVGONLNNAKIFWSSACFOELHCKNSSI 100

QY 95 RFTGTMDAFVKIVHEGSTRMLSGPLATVMTVPAATYATYTAVDQKAFGLRALSXL 154  
 DB 101 KFNCTLEAFTRKIASVEGISTLMRGISLTLMAIPANVYFSGEYIRQD-VSPASTYPTL 159

QY 155 APWAGALALGTVTVISPLEMRTKLQAOHVSRE-----LGACVTRTAQAQGW 204  
 DB 160 NPLECGAIARVAFATSIAPLELVKTRLOSIPRSSKSTKTMWVKDLNFTROEMKVVGPS 219

QY 205 RSLMLGCPALRDVPFSALYMFNEVLVS--WLNGLR--PKDQSYSG--SFVAGGISG 258  
 DB 220 RALFKGLITLMDVDFPSALYMSYELCKERLMDSTRASDAMVHFTNFSAGCISG 279

QY 259 TVAAVLTLPFDVVKTOROVAL-----GAMAVRNPLHVDSTWLLNRIRAESGTRGL 311  
 DB 280 MIAAICTHPEFDVGTWQISMMNNSDPKGNRS-----RNMKFLFTIYRTGLAAL 331

QY 312 FAGFLPRITKAPSCAIMSTYEFKSEFORQRLNQ 345  
 DB 107 ATATYFTTYNLSLWMLKKKCCRAKSPKMTPEWSAAVAIVARTIAVTVVSPTEMI 166

Db 332 YTGIAARVYIKIRPSCAIMISSYEISKVFQNKHLQ 366

RESULT 5

hypotheical protein At2g46320 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: D84901

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L. eus, D.; Nierman, W.C.; White, O.; Elsen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; MUID:20083487

A:Accession: D84901

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-358 <STO>

A:Cross-references: GB:AE002093; NID:96598574; PIDN:AAFL18629.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g46320

A:Map position: 2

Query Match 19.0%; Score 348; DB 2; Length 358;

Best Local Similarity 28.1%; Pred. No. 2,8e-22;

Matches 103; Conservative 61; Mismatches 137; Indels 66; Gaps 10;

QY 13 QOMVASTGAVVTSLEFMPPLDVYKVRLOQSRPMSASELMPSSRLMSLTAKKCLLYC-- 70

Db 23 ERLASAGGAFTISAVVNPDLVAKTRLOAAGVPQY-----GSCRIGCD 68

QY 71 -----NGVLEPLTLCNGARCATWFDPRFRGTMDAFYKIVHEGRTLMGSL 119

Db 69 TNSTLVHDKRNSAPGMCRTISATYCS-----DNQYKGLDVFYKTIHQSTGT----- 118

QY 120 PATLV-----VMTPATVATFYTAVDQ-----LKAFLGRLATSDLYAPVAGALARTGVYI 171

Db 119 --TLCKKGLGFGVGIWPCYDFPNIMEFTTEKSPSLTYVPLVATITARSLACISC 176

QY 172 SPLELMRTKLOAHVSYR--ELGACVRYAV-----AOGGMSLWLGMPALRDVPE 221

Db 177 YVELLARTKQAFKGTQRNVKTLVDVNPVYKGSNNQYRMLMTGLGAQLADVPF 236

QY 222 SALVWNVLYVSKWLN---GLRPKQDTSVGSMPVAGISGTAAVLTLPEDVYKQROVA 278

Db 237 SAICWSTLEPTRRSIQSANGEEPRAGSTIGANFAGFVAGAAVAAATCPDLDAKTRQTE 296

QY 279 LGAMEAVRNPLHVDSTWLLRLIRAESGTLGFLPRTIIRKAPSCAIMISTYEFGRS 338

Db 297 KMTDRAMFM-----TTROTIAEIMRDGMRGMSGAGARVAGSVAIVSVFVYVY 350

QY 339 FFORLNQ 345

Db 351 GLHNFHQ 357

RESULT 6

T00582

probable mitochondrial carrier protein [imported] - Arabidopsis thaliana

N:Alternate names: hypothetical protein T27E13.10

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 23-Mar-2001

C:Accession: T00582; C84705

R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L. eus, D.; Nierman, W.C.; White, O.; Elsen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana.

A:Reference number: AB6141; MUID:21016719

A:Accession: T00582

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-331 <ROU>

A:Cross-references: EMBL:AC004165; NID:93150396; PID:93150404

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y. M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L. eus, D.; Nierman, W.C.; White, O.; Elsen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; MUID:20083487

A:Accession: C84705

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-331 <STO>

A:Cross-references: GB:AE002093; NID:93150404; PIDN:AAFL18629.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g30160; T27E13.10

A:Map position: 2

A:Introns: 263/3

C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

C:Keywords: mitochondrion

Query Match 18.8%; Score 343; DB 2; Length 331;

Best Local Similarity 27.8%; Pred. No. 6,8e-22;

Matches 99; Conservative 58; Mismatches 133; Indels 66; Gaps 7;

QY 3 DDDPAGISPLQD-----MVASGTGAVVTSLEFMPPLDVYKVRLOQSRPMSASELMPSS 54

Db 22 DHPALIVPAQNTTLKFMQWVAGSIAGSVEHMAFEPDVTYKHAQR----- 70

QY 55 RLMSLSYTKMKCLYCNVLEPLTLCNGARCATWFDPRFRGTMDAFYKIVHEGRT 114

Db 71 -----SC-----PIKIGIRQAFRSIIRKIDGESA 94

QY 115 LMSGPLATLVMTPATVATFYTAVDQ-----LKAFLGRLATSDLYAPVAGALARTGVYISPL 174

Db 95 LRGVWAGLGGPMAHYFSTYEVSKFLSG--GNPNNSAAHISGVFATISSDAVFTPM 153

QY 175 ELMRKLOAHVSYRELGAQVTRTAAVAGGMSLWLGMPALRDVPEFSALYFNTELYKS 234

Db 154 DMVKQRLQNGCYGVMDICIRVTRBEQFAFYASRTVLMNPFVHTTYEAVYR 213

QY 235 WINGLRP-----KQDTSVGSMPVAGISGTAAVLTLPEDVYKQROVALAMEAVRN 288

Db 214 GLREMLPEHVAAGEDEGMLIYATGAAAGLAAVTPPLDVYKQLO--CQGVCGCDRK 272

QY 289 PLHVDSTWLLRLIRAESGTLGFLPRTIIRKAPSCAIMISTYEFGRSFOFLN 344

Db 273 SSSISD--VFRTYKKGIRGLARGWLPRLMFHAPAAIICMSTYETVYKSPFDOLN 325

RESULT 7

AB6205

hypotheical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: AB6205

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, ansen, N.F.; Hughes, B.; Huiar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matel, R.; Matzila Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: AB6141; MUID:21016719

A:Accession: AB6205

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-781 <STO>

A:Cross-references: GB:AE005172; NID:96954043; PIDN:AAFL2217.1; GSPDB:GN00141

C:Genetics:



A:Cross-references: EMBL:Z49274; NID:g809585; PIDN:CAA89275.1; PID:g809586; GSPDB:GN0001  
A:Experimental source: strain AB927  
A:Genetics:  
A:Gene: MIPS:YPR021C  
A:Map position: 16R  
C:Superfamily: probable carrier protein YPR021C; ADP, ATP carrier protein repeat homology  
C:Keywords: duplication; transmembrane protein  
F:527-615/Domain: ADP, ATP carrier protein repeat homology <ACPI>  
F:621-711/Domain: ADP, ATP carrier protein repeat homology <ACP2>  
F:724-814/Domain: ADP, ATP carrier protein repeat homology <ACP3>

Query Match 15.6%; Score 284.5; DB 1; Length 902;  
Best Local Similarity 25.8%; Pred. No. 2.4e-16;  
Matches 86; Conservative 53; Mismatches 127; Indels 67; Gaps 6;

16 VASGTGAVVTSLEFMPDLYVVKRLQSORSMASELPPSSRLMSLTWKCKLYCNGVLE 75  
Db IAGCIGATV-----VYPIDIKTRMOQR-SLA----- 565  
Qy PLYLCPNGARCATWFDPTFTGTMDAFVKIVRHGSTRMLSGPLATVMTVPATATFT 135  
Db 566 -----QKNSIDCLIKIISREGIKGLYSGLQPOLIGVAPERRKIKLT 606  
Qy 136 AVDQKAFKLGALTSDLVAPVAGALRLGTVTVISPELMKTKLAQH---VSAREL 191  
Db 607 VVDFMNRRLTDKNKGLSPPEIISGASAGACOVFTNPLEIVKIRLQVSDVGEIIQQA 666  
Qy 192 GACVRAVAGGMRSLMLGMPALRDVPFSALYRNTELVASWNLGRPKQOT-----S 246  
Db 667 NETAQTQVKKLRLGLYNGVAAACLMNDVPSAIYPTTYHLKRLDLPDNDPKTKRNLK 726  
Qy 247 VGMSEVAGGISTGVAVLTLPFDVVTQVQVALGAMAVRNPVLDHDSWMLRLRIAS 306  
Db 727 TWELLTAGAIAMPAPAFLTTPDVIKTRQI-----DRKRETKYNGIFHAIKITKEE 780  
Qy 307 GTKGLPAGFLPRIIKAPSCAIIISTYERKSF 339  
Db 781 SFRSEFKGGARHLRSSPQFGFTLAAVELKGF 813

RESULT 11  
S69050  
probable membrane protein YPL134C - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 22-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 04-Mar-2000  
C:Accession: S69050  
R:Hall, J.; DePaulo, T.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.;  
submitted to the EMBL Data Library, December 1995  
A:Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.  
A:Reference number: S69040  
A:Accession: S69050  
A:Molecule type: DNA  
A:Residues: 1-310 <HAL>  
A:Cross-references: EMBL:U43703; NID:g1244769; PID:g1244780; MIPS:YPL134C  
C:Genetics:  
A:Map position: 16L  
A:Note: YPL134C  
C:Superfamily: ADP, ATP carrier protein; ADP, ATP carrier protein repeat homology  
C:Keywords: duplication; transmembrane protein  
F:8-109/Domain: ADP, ATP carrier protein repeat homology <ACPI>  
F:19-35/Domain: transmembrane \*status predicted <TM>  
F:119-205/Domain: ADP, ATP carrier protein repeat homology <ACP2>  
F:212-301/Domain: ADP, ATP carrier protein repeat homology <ACP3>

Query Match 15.4%; Score 281.5; DB 2; Length 310;  
Best Local Similarity 26.6%; Pred. No. 1.2e-16;  
Matches 93; Conservative 58; Mismatches 125; Indels 73; Gaps 10;

12 LQGVASTGAVVTSLEFMPDLYVVKRLQSORSMASELPPSSRLMSLTWKKC 66  
Db IAGCIGATV-----VYPIDIKTRMOQR-SLA----- 53

Qy 67 LLYCNGVLEPVLICPNGARCATWFDPTFTGTMDAFVKIVRHGSTRMLSGPLATVMT 126  
Db 54 -----AKAAV-----DHYTGMDCLTKIVKREGRSHLYKGTISPLME 91  
Qy 127 VPATATFTAYDQKAFKLGALTSDL-----LYAPVAGALALRGTVTISPLELMR 178  
Db 92 APRRAIKFSGNDTFQFFYKKIPTPENGENTOKAIATISGASAGAVEF-----VAPPELVK 147  
Qy 179 TKLQAGVSVRELGACVRAVAGGMRSLMLGMPALRDVPFSALYV-FNYELVSWLN 237  
Db 148 IRLQDVNSQFKTPIEVKKSVYKGVLSLENGLEATIRHVLNMGYFGLIIRK----- 203  
Qy 238 GLRPDQTSVMS---FVAGGISTGVAVLTLPFDVVKTRQVVALGAMAVRNPLHDS 294  
Db 204 -LLPAKISTEKTNDLGAIGGVGCLLNPFDVVKSRIGRSSQPL---RKYMSLPS 259  
Qy 295 TWLLRIRAESEGTGKGLFAGFLPRIIKAPSCAIIISTYERKSFQRL 343  
Db 260 VLLYVR-----EEGFKALYKGFAPKVKRLAPGGGLLVFTVNMDFREY 304

RESULT 12  
D84798  
probable mitochondrial carrier protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: D84798  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.  
N.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Umayam, L.; Tallon,  
euss, D.; Nierman, W.C.; White, O.; Elsen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487  
A:Accession: D84798  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-348 <STO>  
A:Cross-references: GB:A8002093; NID:g4895195; PIDN:ADJ32782.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g37890  
A:Map position: 2  
C:Superfamily: ADP, ATP carrier protein; ADP, ATP carrier protein repeat homology  
C:Keywords: mitochondrion

Query Match 15.2%; Score 277.5; DB 2; Length 348;  
Best Local Similarity 26.6%; Pred. No. 3e-16;  
Matches 93; Conservative 47; Mismatches 140; Indels 69; Gaps 9;

7 AGISPLQGMVASTGAVVTSLEFMPDLYVVKRLQSORSMASELPPSSRLMSLTWKKC 66  
Db 37 AKLGTFOMLAGGAGALSKCTAPLARLTILFQLQ-GMSEGAVISR----- 83  
Qy 67 LLYCNGVLEPVLICPNGARCATWFDPTFTGTMDAFVKIVRHGSTRMLSGPLATVMT 126  
Db 84 -----PILRREAS-----RIINEBGRFAMWGMLVTVVHR 113  
Qy 127 VPATATFTAYDQKAFKLGALTSDL-----YAPVAGALALRGTVTISPLELMR 178  
Db 114 IPTYAVNFYAEKYNLFFNSNPVVOFTIGNTSGNPIVHFVSGGLAGITAAATVPLDLYR 173  
Qy 179 TKLQAGH--VSARELGACVRAVAGGMRSLMLGMPALRDVPFSALYRNTELVKSWL 236  
Db 174 TRLAQORNAIYYQGIENHFRITICREGLIGLXGLGATLLGVGSLATNFAPAESMKLFW 233  
Qy 237 NGLRPDQTSVMSFVAGGISTVAA-----VLTLPFDVVKTRQVVALGAMAV 285  
Db 224 HSHRPNDSDLV-VSLVSGLAGAVSSTDDKLFDAHNVAATVPLDLYRRMQVE-GAGRA 291  
Qy 286 RVNPLHDSWTLRLRIAESEGTGKGLFAGFLPRIIKAPSCAIIISTYER 334  
Db 292 RYNTGLRGTF---KHIFKSEGFKGYRGLPEYKVVPGVGVFTYD 337



Sat Apr 20 10:03:14 2002

us-09-840-787-19.rpr

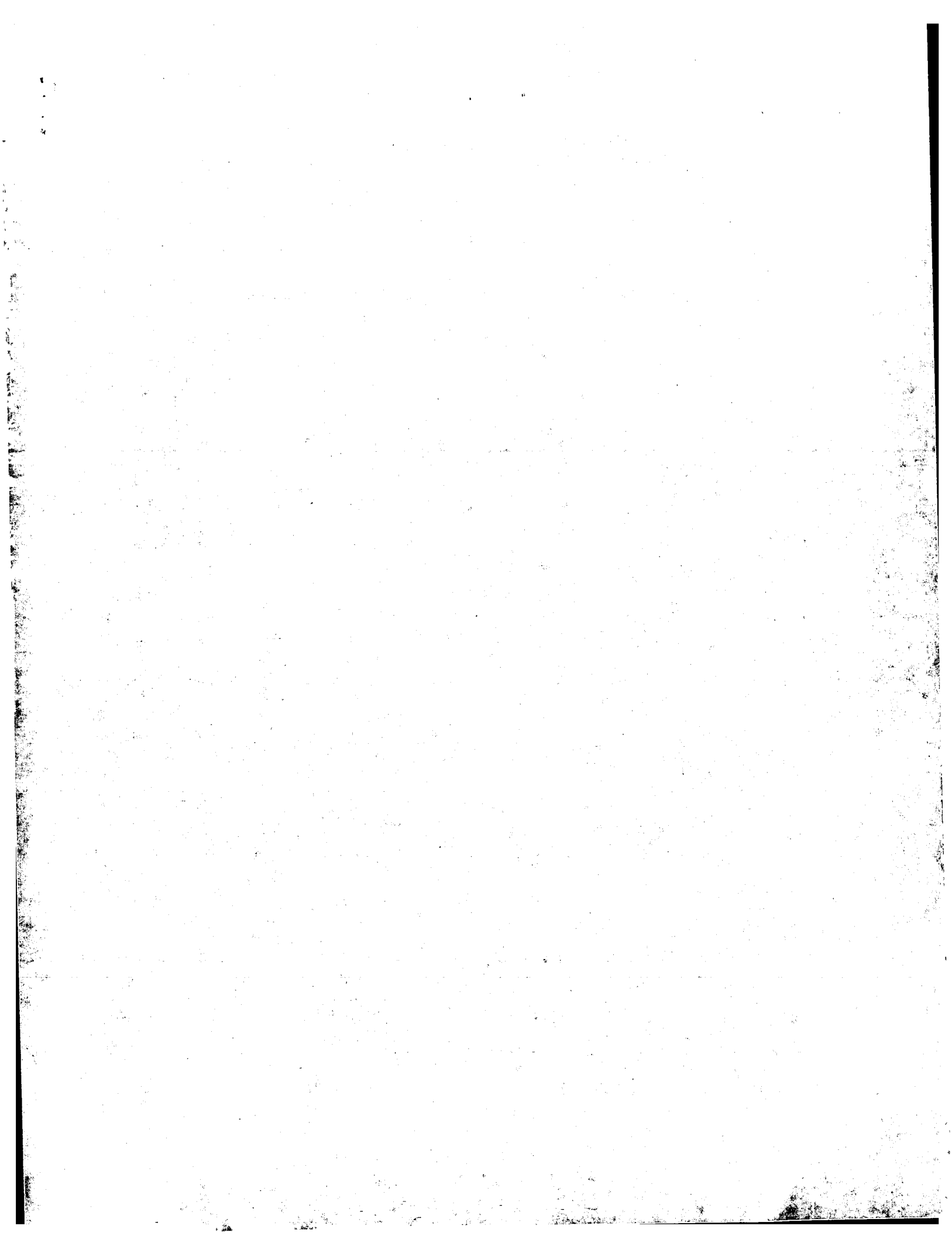
Page 7

```

Db      175  RQLOGRENNPFPOSGIWNYSMLRNAIKTVINEEGFSRLFEFGYKATLARDLPFSALGFAY 234
QY      230  ELVXSMWNLGRPKD---QTSVGMSEFVAGISGTVAAVLTLPEVYNTQ---RQVALGAM 282
        |      |      |      |      |      |      |      |      |      |
Db      235  EKFRQAFRIKQKNGRODELSPNEIILTGACAGGLAITTTTPMDYVYKTRQTOOQPSQSN 294
QY      283  EAVRNPPLHV-----DSTWLLLRIRIAESGTKGLFAGFLPRIIKAPSCAIVMISY 333
        |      |      |      |      |      |      |      |      |      |
Db      295  KSYSTVTHPVYNGRPALNSISLSLTVYQSEGVGLFGESGVGPRFVWISVQSSILPLY 354
QY      334  E 334
        |
Db      355  Q 355

```

Search completed: April 19, 2002, 22:43:22  
Job time: 3355 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2002, 22:42:02 ; Search time 68 Seconds

(without alignments)  
189.255 Million cell updates/sec

Title: US-09-840-787-19

Perfect score: 1828

Sequence: 1 MADQDPAGISPLQOMVASGT.....TYEFGKSFQRLNODRLGG 351

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

## SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	ID	Description
1	648.5	35.5	328 1	Y051_CAEEL Q09461 caenorhabdi
2	474.5	26.0	366 1	Y05F_YEAST P33320 saccharomyc
3	283	15.5	695 1	CMC1_DROME Q9ya73 drosophila
4	282.5	15.5	301 1	MCAT_RAT P97521 rattus norv
5	280	15.3	301 1	MCAT_HUMAN Q43772 homo sapien
6	273	14.9	676 1	CMC2_MOUSE Q9gx44 mus musculu
7	271	14.8	675 1	CMC2_HUMAN Q9aj50 homo sapien
8	268.5	14.7	368 1	Y039_YEAST Q03829 saccharomyc
9	264	14.4	702 1	CMC1_CAEEL Q21153 caenorhabdi
10	260.5	14.3	302 1	Y07L_YEAST Q10248 schizosacch
11	259.5	14.2	307 1	Y07L_YEAST Q09297 saccharomyc
12	255.5	14.0	310 1	UCP2_BRARE Q94720 brachydantio
13	249	13.6	304 1	MRS4_YEAST P23500 saccharomyc
14	245.5	13.4	377 1	RIM2_YEAST P38127 saccharomyc
15	243.5	13.3	322 1	ACR1_YEAST P33303 saccharomyc
16	239	13.1	678 1	CMC1_HUMAN Q75746 homo sapien
17	237	13.0	309 1	UCP2_HUMAN P55851 homo sapien
18	236.5	12.9	310 1	UCP2_CYPCA Q94725 cyprinus ca
19	233	12.7	330 1	GDC_BOVIN Q01888 bos taurus
20	233	12.7	373 1	Y1A6_YEAST P40556 saccharomyc
21	232	12.7	309 1	UCP2_CANFA Q09421 canis famli
22	232	12.6	314 1	MRS3_YEAST P10566 saccharomyc
23	230	12.6	309 1	UCP2_RAT P56500 rattus norv
24	229	12.5	309 1	UCP2_MOUSE P70406 mus musculu
25	227.5	12.4	339 1	ADT_CHLKE P31692 chlorella k
26	226.5	12.4	284 1	PEB8_YEAST P39921 saccharomyc
27	223	12.2	345 1	YDE9_YEAST Q10442 schizosacch
28	221.5	12.1	308 1	UCP3_RAT P56499 rattus norv
29	221	12.1	436 1	BRL_MAIZE P29518 zea mays (m
30	220.5	12.0	323 1	UCP4_HUMAN Q95847 homo sapien
31	219.5	12.0	308 1	UCP3_MOUSE P56501 mus musculu
32	219.5	12.0	357 1	YH62_YEAST P38702 saccharomyc
33	218	11.9	309 1	UCP2_PIG Q97562 sus scrofa

34	213.5	11.7	292 1	AR11_YEAST Q12375 saccharomyc
35	212.5	11.6	313 1	ADT_NEUCR P02723 neurospora
36	210	11.5	307 1	UCP1_HUMAN P25874 homo sapien
37	209	11.4	288 1	UCP1_BOVIN P10861 bos taurus
38	206.5	11.3	311 1	UCP3_BOVIN Q77792 bos taurus
39	205	11.2	318 1	ADT2_YEAST P18239 saccharomyc
40	204	11.2	332 1	GDC_HUMAN P16260 homo sapien
41	204	11.2	305 1	ADT_KLULA P49382 kluyveromyc
42	202.5	11.1	308 1	UCP3_PIG Q97649 sus scrofa
43	201	11.0	311 1	UCP3_CANFA Q9n219 canis famli
44	199.5	10.9	298 1	ADT3_BOVIN P32007 bos taurus
45	199	10.9	335 1	YEA6_YEAST P39953 saccharomyc

## ALIGNMENTS

RESULT	ID	Y051_CAEEL	STANDARD	PRT	328 AA.
AC	Q09461				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	01-NOV-1997	(Rel. 35, Last annotation update)			
DE	POTATIVE MITOCHONDRIAL CARRIER C16C10.1.				
GN	C16C10.1				
OS	Caenorhabditis elegans.				
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;				
CC	Nematoda; Peloderidae; Caenorhabditis.				
OK	NCBI_TaxID=6239;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=BRISTOL N2;				
RA	Lloyd C.;				
RL	Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.				
CC	-1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.				
CC	-1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
DR	EMBL: Z46787; CA86739.1; -				
DR	WormPep; C16C10.1; CE01489.				
DR	InterPro; IPR001993; Mitoch_carrier.				
DR	Pfam; PF00153; mltc_carr; 3.				
DR	PROSITE; PS00215; MITOCH_CARRIER; 1.				
KW	Hypothetical protein; Mitochondrion; Repeat; Transmembrane;				
KW	Transport.				
FT	TRANSMEM 93 113				POTENTIAL.
FT	TRANSMEM 146 166				POTENTIAL.
FT	TRANSMEM 237 257				POTENTIAL.
FT	SEQUENCE 328 AA; 36743 MW; EA1E9E329A764DF6 CRC64;				
SQ					
Query Match		35.5%	Score 648.5;	DB 1;	Length 328;
Best Local Similarity		39.4%	Pred. No. 2.7e-48;		
Matches 136;	Conservative 63;	Mismatches 113;	Indels 33;	Gaps 6;	
OY	10 SPLQOMVAGTAVVTSLEMPPLDVKYRLOBO-PPSMASELMPSRLMSLTWKCLL 68				
DB	3 SPCESGKTTNCSGAPSSSCVPLDVKYRLOBOQTRP-----FPKGECPY 46				
OY	69 YCNGLEPLYLCPNGARATWFOPTRTGTDAFVKIVRHGRTLTWSGLPATLVMTVP 128				
DB	47 YHNGLEMHVAVCEVRCQCEWYQRCNGFRTADLVKIAHKGISLWSGLSPTVMALP 106				
OY	129 ATATFTAYDQAKFL-----CGRALTSDLVAP-----WAGALARIGTVIVIPLELM 177				
DB	107 ATVFYFTTYDNLISVWLKKKMCRCRAFSPEKWPDPMSAAVAGIVARTIATVVSPIEMI 166				



RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pauleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard D.R., Paoletti J.M.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
CC -1- FUNCTION: CALCIUM-DEPENDENT MITOCHONDRIAL SOLUTE CARRIER (BY  
CC SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL  
CC INNER MEMBRANE (BY SIMILARITY).  
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1 (SHOWN HERE), 2 AND 3; MAY BE  
CC PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. ARALAR  
CC SUBFAMILY.  
CC -1- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Y18197. CAB62169.1. -;  
DR EMBL: AE003774.1. AAF57048.1. -;  
DR EMBL: AE003774.1. AAF57049.1. -;  
DR EMBL: AE003774.1. AAF57050.2. -;  
DR -FlyBase: FBgn028646; atalari.  
DR InterPro: IPR002048; EF-hand.  
DR InterPro: IPR002067; Mito\_carrier.  
DR InterPro: IPR001993; Mitoch\_carrifer.  
DR Pfam: PF00036; efhand; 3.  
DR Pfam: PF00153; mito\_carr; 3.  
DR PRINTS: PR00926; MITOCARRIER.  
DR SMART: SM00054; EFh; 3.  
DR PROSITE: PS00018; EF-HAND; 1.  
DR PROSITE: PS00215; MITOCH\_CARRIER; 3.  
KW Calcium-binding; Alternative splicing.  
KW Calcium-binding; Alternative splicing.  
KW Transmembrane; Transport;  
KW Transmembrane; Transport;  
FT TRANSMEM 346 363 1 (POTENTIAL).  
FT TRANSMEM 406 425 2 (POTENTIAL).  
FT TRANSMEM 449 462 3 (POTENTIAL).  
FT TRANSMEM 498 517 4 (POTENTIAL).  
FT TRANSMEM 537 554 5 (POTENTIAL).  
FT TRANSMEM 594 613 6 (POTENTIAL).  
FT CA\_BIND 84 95 EF-HAND 1.  
FT CA\_BIND 118 129 EF-HAND 2.  
FT CA\_BIND 157 168 ANCESTRAL CALCIUM SITE 3.  
FT CA\_BIND 189 200 EF-HAND 4.  
FT DOMAIN 225 237 3 APPROXIMATE TANDEM REPEATS.  
FT REPEAT 225 434 1.  
FT REPEAT 435 524 2.  
FT REPEAT 525 627 3.  
FT VASPLIC 1 52 MISSING (IN ISOFORM 3).  
FT VASPLIC 3 22 MISPPFWMTLVARQOE -> LITSLPN (IN  
FT ISOFORM 2).  
FT GA -> AP (IN REF. 1).  
FT V -> A (IN REF. 1).  
FT V -> L (IN REF. 1).  
SQ SEQUENCE 695 AA; 76753 MW; 8EC93D92031F5B77 CMC64;

Query Match 15.5%; Score 283; DB 1; Length 695;  
Best Local Similarity 27.3%; Pred. No. 1,2e-16;  
Matches 97; Conservative 55; Mismatches 129; Indels 74; Gaps 12;  
OY 2 ADQDPAGISPLQOMVAGS-----TGAVTSLEMTPLDVVKVRLQSGRPSMASLMS 53  
DB 326 AVESPADRSFAQVLESRYFTLGSFAGAVGATV-VYPIDLVKTRMQNR-----AG 376  
OY 54 SLMNLSY-TRKCKLLCYNGVLEPLICNGARCATWEDPFRFTGTMDAPFKYRHET 112  
DB 377 SYIGEVAVRNSWDC-----FRKVVYRHEGF 400  
OY 113 RLTMGLPPLTIFWTPATATYFTAYDQKAFICGRALTSDLVAPVAGALRLGTYVIS 172  
DB 401 MCLYKGLPLQMGVAPERAKITVDLVNDKDKGNIPYAEVLAGGCACASQVFTN 460  
OY 173 PLEMRKTLQ-AQHVSYRELACVR--TAVAGGMRSLWLGMPALRDVPSALWPNY 229  
DB 461 PLEIKIRIQVAGEIA---SGSKIRAMSVRELGLFGYKGRACLLRDVPSALYFTY 517  
OY 230 ELVKSWLNGLRKQDTSVGMSTV-AGISGTYAAVLTLPDVVKYQROVALCAMEAVRN 288  
DB 518 AHTKAM---MADKDYVNPDLTLAAGALAGVPAASIVPADVIKTRLOV-----VARSG 568  
OY 289 PLHVDSTWLLLRIRASGKGLFAGFLPRIIKAPSCAIMSTYEFKSFQRL 343  
DB 569 QTTTGVWDATKKIMAESEPRFMTARVRSRQSGVTLVTE---LIQRL 619  
RESULT 4  
MCAT\_RAT  
ID MCAT\_RAT STANDARD; PRT; 301 AA.  
AC P97521.  
DI 01-NOV-1997 (Rel. 35, Created)  
DI 01-NOV-1997 (Rel. 35, Last sequence update)  
DI 30-MAY-2000 (Rel. 39, Last annotation update)  
DE MITOCHONDRIAL CARNITINE/ACYLCARNITINE CARRIER PROTEIN  
DE (CARNITINE/ACYLCARNITINE TRANSLOCASE) (CAC).  
GN CACT.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RP [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Liver;  
RX MEDLINE=97184657; PubMed=9032458;  
RA Indiveri C., Jacobazzi V., Giangregorio N., Palmieri F.;  
RT "The mitochondrial carnitine carrier protein: cDNA cloning, primary  
RT structure and comparison with other mitochondrial transport  
RT proteins.";  
RL Biochem. J. 321:713-719(1997).  
CC -1- FUNCTION: TRANSPORT OF CARNITINE OR ACYLCARNITINE FROM THE CYTOSOL  
CC TO THE MITOCHONDRIAL MATRIX.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL  
CC INNER MEMBRANE.  
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.  
CC -1- PTR: THE N-TERMINAL IS BLOCKED.  
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X97831. CAA66410.1. -;  
DR InterPro: IPR001993; Mitoch\_carrifer.  
DR Pfam: PF00153; mito\_carr; 3.  
DR PROSITE: PS00215; MITOCH\_CARRIER; 3.





[illegible]

```

Db      526 AGAIAQMPASALTVPADVIKTRLQVAAARGOQTYSVID-----CFRKILREGRPALM 579
OY      313 AGLPLRIITKAAPSCAIMISTYERKSAFF 340
       | : ::::| : || : :|
Db      580 KGAGARVERSSPQFGVTLLTYELLQRWF 607

RESULT      8
YM39_YEAST
ID    YM39_YEAST      STANDARD;          PRT;           368 AA.
AC    003829;
DT     01-NOV-1997 (Rel. 35, Created)
DD     01-NOV-1997 (Rel. 35, Last sequence update)
DR     01-NOV-1997 (Rel. 35, Last annotation update)
DE     PUTATIVE MITOCHONDRIAL CARRIER YMR166C.
GN     YMR166C OR YM8520.15C.
OS     Saccharomyces cerevisiae (Baker's yeast).
OC     Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC     Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CB     NCBI_TaxID=4932;
[1]
RC     SEQUENCE FROM N.A.
RP     STRAIN=S288C / AB972;
RA     Hunt S., Bowman S., Barrell B.G., Rajandream M.A.;
RL     Submitted (May-1995) to the EMBL/GenBank/DBJ databases.
CC     -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
        INNER MEMBRANE (POTENTIAL).
        -I- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
        -I- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC-----
CC     This SWISS-PROT entry is copyright. It is produced through a collaboration
CC     between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC     the European Bioinformatics Institute. There are no restrictions on its
CC     use by non-profit institutions as long as its content is in no way
CC     modified and this statement is not removed. Usage by and for commercial
CC     entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC     or send an email to license@isb-sdb.ch).
CC-----
DR     EMBL; Z49705; CAA89802.1; -.
DR     SGD; S0004776; YMR166C.
DR     InterPro; IPR001993; Mitoch_carrler.
DR     Pfam; PF00153; mito_carr; 3.
DR     ProSITE; PS00215; MITOCH_CARRIER; 2.
KW     Hypothetical protein; Mitochondrion; Inner membrane; Repeat;
KM     Transmembrane; Transport.
FT     TRANSMEM 110 130      POTENTIAL.
FT     TRANSMEM 152 172      POTENTIAL.
FT     TRANSMEM 260 280      POTENTIAL.
SQ     SEQUENCE   368 AA; 40992 MW; B583100018DF045D CRC64;

Query Match      14.7%, Score 268.5; DB 1; Length 368;
Best Local Similarity 24.9%; Pred. No. le-15;
Matches 90; Conservative 56; Mismatches 136; Indels 79; Gaps 9;

OY      2 ADOPAGISPLQOMVAS-GTGAVVYSLFMPPLDVKYVRLQSGRPSMASBLMPSSRWLSLS 60
       :| :| :|| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      46 SDED--LSPIMHCYVSSGISGKIGDSAMHSILDYKTKQQG----- 83

OY      61 YTKMKCLLYCGNVLEPIYLCPNGARCATWFODPRTRETGMADFVKIVNHEGTRT-LMSG 119
       :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      84 -----ADN-----VKKYRNMSIAYRTIMEDEGVRRGLYGGY 114

OY      120 PATLVMTVPATAIYFTYADOLKAFPLCGALSLSLTAPYNAVAGALARLGIVYYISPFLMRT 179
       :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      115 MAAMLGSPSPAIFEGFYETRYKRRTIEDMWINDITTHLSAGLFDEFISSFYYPSEVEIKT 174

OY      KIQAQ-----HVSYRELGA CVRTAAVAQGWSRSLMGCTALRDVPFSALWYNT 229
       :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      175 RLQLOGRRNNFPQSGYNYSNLNRNAIKTVIKEEGFRSLPFQYKATKLANDLPFSLQGFAY 234

OY      230 ELVXSWLNGLRPKD---QTSVGMSFVAQGISGTVAVALLPEDVVKTQ---ROYALGAM 282
       | : :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

```

```

Db 235 EKEROLAKIEQKGRDGEFLSPNEILTGACAGLAGIITTPMDVYKTRVOTQOPPSQSN 294
Qy 283 EAVRNPLHV-----DSTWLLLRIRAESGKGLFAGFLPRIKAPSCAIMISRY 333
Db 295 KSYSTHPRHVTVNGRPALNSISLSIRYVQSEGLVGFSSGVPFRVMTSVOSIMLLY 354
Qy 334 E 334
Db 355 Q 355

RESULT 9
CMC1_CAEEL STANDARD: PRT: 702 AA.
ID CMC1_CAEEL
AC 021153;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE CALCIUM-BINDING MITOCHONDRIAL CARRIER K02F3.2.
GN K02F3.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCB1_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Bentley D.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CALCIUM-DEPENDENT MITOCHONDRIAL SOLUTE CARRIER (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY. ARAJAR
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U00052; AKK21421.1; -
DR Wormpep; K02F3.2; GE01348.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam; PF00153; efhand; 3.
DR PRINTS; PR00926; mito_carr; 3.
DR PRINTS; PR00927; ADPTRNSLCASE.
DR SMART; SM00054; EFh; 3.
DR PROSITE; PS00018; EF_HAND; 2.
DR PROSITE; PS00215; MITOCH_CARRIER; 3.
KW Hypothetical protein: Mitochondrion; Inner membrane; Repeat;
KW Transmembrane; Transport; Calcium-binding.
FT TRANSMEM 368 385
FT TRANSMEM 429 448
FT TRANSMEM 471 484
FT TRANSMEM 521 540
FT TRANSMEM 560 577
FT TRANSMEM 617 636
FT CA_BIND 106 117
FT CA_BIND 140 151
FT CA_BIND 179 190
FT DOMAIN 211 222
FT DOMAIN 247 252
FT REPEAT 247 252
FT REPEAT 457 457
FT REPEAT 548 548
FT REPEAT 650 650
SQ SEQUENCE 702 AA; 78170 MW; 41AB9EF9A1F018D0 CRC64;

```

```

Query Match. 14.4%; Score 264; DB 1; Length 702;
Best Local Similarity 27.3%; Pred. No. 5, 1e-15;
Matches 91; Conservative 46; Mismatches 128; Indels 68; Gaps 9;

Qy 16 VASGTAVVTSLEMTPLDVVKVRLQSORPSMASSELPSSRLMSLSYTKWICLLYGVLE 75
Db 373 VAGACGA--TANY--PIDLVKTRMGNKRTSS----- 400
Qy 76 PLYLCPNGARCATWFDQDPTRETGTMDAFAVKYIRHIGSTRTLMSGPLATLVMPATAIYFT 135
Db 401 -----FVGEVMYKNSIDCEKRYKVFEGLLGLGLPQLYGVAPERAKIKLT 446
Qy 136 AVDOAKFLCGRALTS-----LYAPMVAGALARLTGVVISLELMRKQLQ AQHVSTRE 190
Db 447 MNDYMR-----DKFRKDKIPILYGEIILAGTGMCOVFTNLEIVKIRLQTAGEVQAG 501
Qy 191 LGACVRTVAAGGWSRLMGWGPTRALRDVPFSALYFNVYELKWSMLNGLRPKDRTSVGMS 250
Db 502 KKIGFVYLKELGFLGKGRACFLRDIPIFAIYFPAVAAHK--LASNDEGNMSPGTL 559
Qy 251 FVAGISGTVAAVLTLPEDVYKTRQOVAAGAEAVRNPLHVDSTWLLLRIRAESGKGL 310
Db 560 FASAFIAGVPAGLVTPADVITKTRQV-----AARAGQTYNGVIDCARLKKEGPM 613
Qy 311 LFAGFLPRIKAPSCAIMISYERKSEFQRL 343
Db 614 LMKGTAAVCRSSPQFAVTLTYE---VLQRL 642

RESULT 10
YDIK_SCHPO STANDARD: PRT: 302 AA.
ID YDIK_SCHPO
AC Q10248;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PURNATIVE MITOCHONDRIAL CARRIER C4G9.20C.
GN SPAC4G9.20C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCB1_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (POTENTIAL).
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z69727; CA93570.1; -
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam; PF00153; mito_carr; 3.
DR PROSITE; PS00215; MITOCH_CARRIER; 2.
KW Hypothetical protein: Mitochondrion; Inner membrane; Repeat;
KW Transmembrane; Transport.
FT TRANSMEM 9 29
FT TRANSMEM 71 91
FT TRANSMEM 116 136
FT TRANSMEM 216 236

```



Query Match	14.08;	Score 255.5;	DB 1;	Length 310;
Best Local Similarity	24.68;	Pred. NO. 1.1e-14;		
Matches 83;	Conservative 44;	Mismatches 135;	Indels 75;	Gaps 7;

RESULT	13
MSR4_YEAST	
ID	MSR4_YEAST
AC	P23500;
DT	01-NOV-1991 (Rel. 20, Created)
DT	01-NOV-1991 (Rel. 20, Last sequence update)
DT	15-DEC-1996 (Rel. 37, Last annotation update)
DE	MITOCHONDRIAL RNA SPLICING PROTEIN MSR4.
GN	MSR4 OR YKR052C.
OS	Saccharomyces cerevisiae (Baker's yeast).
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX	NCBI_TaxID:4932;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN:NM1301.
RX	MEDLINE-9110815; PubMed-1703236;
RA	Wiesenberger G., Link T.A., von Ahesen U., Walther M., Schweyen R.J.,
RT	MR3 and MSR4, two suppressors of mRNA splicing defects in yeast,
TI	are new members of the mitochondrial carrier family."

DR	SOD	\$0001760;	MRS4.
DR	InterPro:	IPR00193;	Mitoch.carrier.
DR	Pfam:	PF00153;	mito_carr; 3.
DR	PROSITE:	PS00215;	MITOCH.CARRIER; 2.
KW	Mitochondrion:	Inner membrane;	Repeat; Transmembrane; Transport.
KW	MRNA processing.		
FT	REPEAT	9	110
FT	REPEAT	111	202
FT	REPEAT	203	304
FT	TRANSMEM	23	41
FT	TRANSMEM	83	102
FT	TRANSMEM	120	139
FT	TRANSMEM	175	194
FT	TRANSMEM	209	228
FT	TRANSMEM	275	288
SO	SEQUENCE	304 AA;	33307 MW; 5ABBF985B8547EDC CRC64;

Db	21	APLHSQLLAGFAGMEHSLMEFPIDALTRVOAGLINKAAS----	61
QY	69	YCNGBLEBLYLCPNARGCATWEDPTEPTGMDAFVKIVRHGRTMLSGLPATLYMTPV	122
Db	62	-----TGHSIQISKISTMEGSMALWKGVQSVILGAP	93
QY	129	ATATYFTYDOLKAFPLCGRALTSDL--YAPM--VAGALARTGYTVISPLMLRTKLOA	18
Db	94	AHAAYFFGYECCAKRLIS---PEDMQTHQPKMTALSGTIATIAADALMNPDPYKORQL	150
QY	184	QHVASRETGACVPNVAHQGGWRSIMLGMGPALNDVPPSALYXNRYNELVKSMLGRLPKD	244
Db	151	D--TNLRKWNANTKQIYONEGFPAAYYSIPTLANNIPPAALNFMYSASKFFN--PON	205
QY	244	QTSVGMSEFVAGGISQTYAAAVLTLPEDVYKOTQROYALGMEAVRNPPLHVDST--LLLR	302
Db	206	SYNPLIHOLCGISGARCALTTPLDCCIKTYLQVR--GSEVLSLEIKKDANTFORASRAI	263
QY	303	RAESGTGLPAGFLPRITIKAPSCALIMISTYEFKSEFFOR	342
Db	264	LEVHGKMGFWGRLKPRIVANIPALVIMWTLYECCAKHPLMK	303

```

RESULT 14
RIM2_YEAST STANDARD; PRT; 377 AA.
AC P38127;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE MITOCHONDRIAL CARRIER PROTEIN RIM2.
GN RIM2 OR YBR192W OR YBR1402.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=93348777; PubMed=8346681;
RA Demolis N., Mallet L., Bussereau F., Jacquet M.;
RT "RIM2, MS1 and PGI1 are located within an 8 kb segment of
RT Saccharomyces cerevisiae chromosome II, which also contains the
RT putative ribosomal gene L21 and a new putative essential gene with a
RT leucine zipper motif."
RT Yeast 9:645-659(1993).
RL
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z21487; CAA9678.1; -
CC EMBL; Z36061; CAA85154.1; -
CC PIR; S36081; S36081.
CC SGD; S0000396; RIM2.
CC InterPro: IPR001993; Mitoch_carrier.
CC Pfam: PF00153; mito_carrier_3.
CC PROSITE: PS00215; MITOCH_CARRIER; 1.
CC Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
CC KW SEQUENCE 377 AA; 42101 MW; A3C61383C3B3743B CRC64;
SO
Query Match 13.4%; Score 245.5; DB 1; Length 377;
Best Local Similarity 22.2%; Pred. No. 9.5e-14; Indels 109; Gaps 11;
Matches 84; Conservative 58; Mismatches 127;

```

```

OY 317 PRIKAPSCAISMISTYE 334
|:::|:::|:::|
DB 353 PHMKRTVPNSIMIGTWE 370

RESULT 15
ACR1_YEAST STANDARD; PRT; 322 AA.
AC P33303;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE REGULATOR OF ACETYL-COA SYNTHETASE ACTIVITY.
GN ACR1 OR YJR095W OR J1921.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=94203187; PubMed=7908717;
RX Fernandez M., Fernandez E., Rodicio R.;
RA "ACR1, a gene encoding a protein related to mitochondrial carriers,
RT is essential for acetyl-CoA synthetase activity in Saccharomyces
RT cerevisiae."
RT Mol. Gen. Genet. 242:727-735(1994).
RL
CC [2]
CC SEQUENCE FROM N.A.
CC Ramezani Rad M., Kirchhath L., Hollenberg C.P.;
CC Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ESSENTIAL FOR ACETYL-COA SYNTHETASE ACTIVITY. ESSENTIAL
CC FOR GROWTH ON ETHANOL AND ACETATE AS SOLE CARBON SOURCES.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC OR PEROXISOMAL.
CC -1- INDUCTION: BY ETHANOL OR ACETATE AS SOLE CARBON SOURCES. REPRESSED
CC BY GLUCOSE.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z25485; CAA80973.1; -
CC EMBL; Z49595; CAA89624.1; -
CC PIR; S36407; S36407.
CC PIR; S43280; S43280.
CC SGD; S0003856; SFC1.
CC InterPro: IPR001993; Mitoch_carrier.
CC Pfam: PF00153; mito_carrier_3.
CC PROSITE: PS00215; MITOCH_CARRIER; 3.
CC Transmembrane; Transport; Repeat.
CC KW YREH1ENLGIETKNDTPKPKPLK -> RKGAFOKIWYSRR
CC FT CONFLICT 300 322 MTHOSOSH (IN REF. 1).
CC SO SEQUENCE 322 AA; 35340 MW; EC29718A0F5011A7 CRC64;
SO
Query Match 13.3%; Score 243.5; DB 1; Length 322;
Best Local Similarity 22.1%; Pred. No. 1.2e-13;
Matches 81; Conservative 55; Mismatches 129; Indels 101; Gaps 9;

```

'Sat Apr 20 10:03:14 2002

us-09-840-787-19.rsp

Page 11

```

OY 127 VPATVYTTAVDQAKAPL-----CGRALTSIDJAPMVAAGALARIQTVIVISPELMTKQO 1822
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 83 IPKMAIRFSSFEYRTLLVYNKESIVSTGNTFVAGAGITE--AYLVNPMVEVWKIRO 140
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 183 AOHVSYRELGA-----CVRTAVAGOGKRSJLMLGMPALTALRDVFPFSAJYMFNELYK 233
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 AQHLLPSEPNAGPKYNNAIHAAYITVKEEVSALYRGVSLTAARQAINQANCNTVYSKUL 200
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 234 SWLNGLRKQDTSYGMSTVAGISGTYAAAVLTLPEDVYKTKORQ-----V 277
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 201 EFLONYHOMDVLPSWETSJGLISGAIGPSPMAJLDIKTRILQDKSISLEKSGMKKII 260
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 278 ALGAEAVRVNPLVHVDSTWLLLRIRIRAESGTKGLFAGLEPRIITKAAPSCAII MSTYEFG 357
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 261 TIGA-----QLKEEGFRALYKTIPTFRVNRVAPAGCAVTTYVIEYR 301
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 338 SFFQRL 343
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 302 EHLENTL 307
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: April 19, 2002, 22:52:36  
Job time: 634 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2002, 21:46:02 ; Search time 62.21 Seconds  
(without alignments)  
126.968 Million cell updates/sec

Title: US-09-840-787-19

Perfect score: 1828

Sequence: 1 MADQDPAGISPILOQWASGT.....TYEFGKSFQRLNDRLLG 351

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5  
212252 segs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/1aa/PTOUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/1aa/Backfilest.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1828	100.0	351	2	US-08-933-750C-19
2	1828	100.0	351	4	US-09-234-613-19
3	324	17.7	328	4	US-09-068-140A-15
4	320	17.5	289	1	US-09-068-140A-10
5	237	13.0	299	1	US-08-518-878B-56
6	237	13.0	299	2	US-08-470-868A-56
7	237	13.0	309	1	US-08-518-878B-51
8	237	13.0	309	2	US-08-807-861A-51
9	237	13.0	309	3	US-08-470-868A-51
10	237	13.0	309	3	US-09-210-681-51
11	237	13.0	309	3	US-08-946-719A-51
12	236	12.9	320	2	US-08-933-750C-12
13	236	12.9	320	4	US-09-234-613-12
14	233.5	12.8	469	4	US-09-188-930-39
15	221.5	12.1	308	2	US-08-937-466-2
16	221.5	12.1	308	3	US-09-172-528-2
17	221.5	12.1	308	3	US-09-318-199-2
18	221.5	12.1	308	4	US-09-503-579-2
19	219.5	12.0	432	2	US-08-937-466-4
20	219.5	12.0	432	2	US-09-172-528-4
21	219.5	12.0	432	3	US-09-318-199-4
22	219.5	12.0	432	3	US-09-503-579-4
23	200	10.9	303	1	US-08-518-878B-37
24	200	10.9	303	1	US-08-294-522B-36
25	200	10.9	303	2	US-08-807-861A-37
26	200	10.9	303	2	US-08-470-868A-37
27	200	10.9	303	3	US-09-210-681-37

28	200	10.9	303	3	US-08-946-719A-37
29	198.5	10.9	306	5	PCT-US94-09799-1
30	197.5	10.8	312	4	US-09-142-565-2
31	194	10.6	307	2	US-08-807-861A-56
32	194	10.6	307	3	US-09-210-681-56
33	194	10.6	307	3	US-08-946-719A-56
34	192.5	10.5	311	2	US-08-775-009-32
35	185.5	10.1	298	3	US-08-961-871-10
36	181	9.9	256	2	US-08-937-466-6
37	181	9.9	256	2	US-09-172-528-6
38	181	9.9	256	3	US-09-318-199-6
39	181	9.9	256	4	US-09-503-579-6
40	177	9.7	311	2	US-08-775-009-33
41	125	6.8	312	4	US-09-188-930-142
42	90.5	5.0	3815	4	US-09-428-517-3
43	87	4.8	998	2	US-08-449-645A-20
44	87	4.8	998	2	US-08-702-367A-20
45	87	4.8	998	5	PCT-US95-04681-20

## ALIGNMENTS

RESULT 1  
US-08-933-750C-19  
Sequence 19, Application US/08933750C  
Patent No. 5932442  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Shah, Purvi  
APPLICANT: Au-Young, Janice  
APPLICANT: Yue, Henry  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/933 750C  
FILING DATE: September 23, 1997  
CLASSIFICATION: 536  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 351 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: SYNCOAT01

Sequence 37, Appl
Sequence 1, Appl
Sequence 2, Appl
Sequence 56, Appl
Sequence 56, Appl
Sequence 32, Appl
Sequence 10, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 33, Appl
Sequence 142, App
Sequence 3, Appl
Sequence 20, Appl
Sequence 20, Appl
Sequence 20, Appl

CLONE: 724157  
US-08-933-750C-19

Query Match 100.0%; Score 1828; DB 2; Length 351;  
Best Local Similarity 100.0%; Pred. No. 1,1e-179;  
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MADOPAGISPLQOMVASGTGAIVTSLFMTPLDVYKVRLOSQRPSMASLMPSSRLMSLS 60  
DB 1 MADOPAGISPLQOMVASGTGAIVTSLFMTPLDVYKVRLOSQRPSMASLMPSSRLMSLS 60  
QY 61 YTKMKCLYCNGLVLEPLYLCPNGARCATWFOPTFTGTMDAFVKIVRHEGRTTMSGLP 120  
DB 61 YTKMKCLYCNGLVLEPLYLCPNGARCATWFOPTFTGTMDAFVKIVRHEGRTTMSGLP 120  
QY 121 ATLVTVPATATFTAYDOLKAFLCGRALTSPLYAPMAGALARIGTVVISPLELMTK 180  
DB 121 ATLVTVPATATFTAYDOLKAFLCGRALTSPLYAPMAGALARIGTVVISPLELMTK 180  
QY 181 LOAHVSYRELACVRYTAAOGGWSLMLGMPALRDVPFSALYWFNTELYKSWLNGLR 240  
DB 181 LOAHVSYRELACVRYTAAOGGWSLMLGMPALRDVPFSALYWFNTELYKSWLNGLR 240  
QY 241 PKDQTSVGMSFYAGISGTVAAVLTLPDVYKTOBOVALGAEAVRVNPLHVDSTWLLR 300  
DB 241 PKDQTSVGMSFYAGISGTVAAVLTLPDVYKTOBOVALGAEAVRVNPLHVDSTWLLR 300  
QY 301 RIRASSTKGLFAGFLPRITKAAPSCAIMISTYERKSFQRLNDRLLGG 351  
DB 301 RIRASSTKGLFAGFLPRITKAAPSCAIMISTYERKSFQRLNDRLLGG 351

## RESULT 2

US-09-234-613-19  
Sequence 19, Application US/09234613  
Patent No. 6132973

## GENERAL INFORMATION:

APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Shah, Purvi  
APPLICANT: Au-Yang, Janice  
APPLICANT: Yue, Henry  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/234,613  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/933,750  
FILING DATE: September 23, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PR-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 351 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: SYNCOAT01  
CLONE: 724157

US-09-234-613-19

Query Match 100.0%; Score 1828; DB 4; Length 351;  
Best Local Similarity 100.0%; Pred. No. 1,1e-179;  
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MADOPAGISPLQOMVASGTGAIVTSLFMTPLDVYKVRLOSQRPSMASLMPSSRLMSLS 60  
DB 1 MADOPAGISPLQOMVASGTGAIVTSLFMTPLDVYKVRLOSQRPSMASLMPSSRLMSLS 60  
QY 61 YTKMKCLYCNGLVLEPLYLCPNGARCATWFOPTFTGTMDAFVKIVRHEGRTTMSGLP 120  
DB 61 YTKMKCLYCNGLVLEPLYLCPNGARCATWFOPTFTGTMDAFVKIVRHEGRTTMSGLP 120  
QY 121 ATLVTVPATATFTAYDOLKAFLCGRALTSPLYAPMAGALARIGTVVISPLELMTK 180  
DB 121 ATLVTVPATATFTAYDOLKAFLCGRALTSPLYAPMAGALARIGTVVISPLELMTK 180  
QY 181 LOAHVSYRELACVRYTAAOGGWSLMLGMPALRDVPFSALYWFNTELYKSWLNGLR 240  
DB 181 LOAHVSYRELACVRYTAAOGGWSLMLGMPALRDVPFSALYWFNTELYKSWLNGLR 240  
QY 241 PKDQTSVGMSFYAGISGTVAAVLTLPDVYKTOBOVALGAEAVRVNPLHVDSTWLLR 300  
DB 241 PKDQTSVGMSFYAGISGTVAAVLTLPDVYKTOBOVALGAEAVRVNPLHVDSTWLLR 300  
QY 301 RIRASSTKGLFAGFLPRITKAAPSCAIMISTYERKSFQRLNDRLLGG 351  
DB 301 RIRASSTKGLFAGFLPRITKAAPSCAIMISTYERKSFQRLNDRLLGG 351

## RESULT 3

US-09-068-140A-15  
Sequence 15, Application US/09068140A  
Patent No. 6281409

## GENERAL INFORMATION:

APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor  
APPLICANT: and Rex Michael Brennan  
TITLE OF INVENTION: Blackcurrant Promoters and Genes  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/068,140A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP96/04807  
FILING DATE: No. 6281409ember 4, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Dinner, Data L.

REGISTRATION NUMBER: 33,680  
REFERENCE/DOCKET NUMBER: C70237  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5017  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 328 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Rides nigrum  
STRAIN: Ben Alder  
US-09-068-140A-15

Query Match 17.7%; Score 324; DB 4; Length 328;  
Best Local Similarity 27.0%; Pred. No. 3.5e-25;  
Matches 92; Conservative 65; Mismatches 128; Indels 56; Gaps 8;

OY 8 GISPLQWVAGSTGAVVTSLEFPTPLDVVKVRLQSORPSMASLMSRLMSLYTKMKCL 67  
DB 33 GLQFQFMIAAGSIAGSIEHMAHYVDLTKRIQ----- 65  
OY 68 LYCNGLVEPLYLCNGARCAWFOPTREFTGMDAFKIVRHEGTRTLMGSLPATLVMTV 127  
DB 66 -----GSCS-----AQSAGLRQALGSLIKKEGPAGLYRGIGAMGLGAG 105  
OY 128 PATATYFAYDQK-AFLCGRALTSDLAPMYAGALRLGTYTVISPLEMRTKIQAOHV 186  
DB 106 PAHAAYFVYEVCKEFTFSGHGPSNSGAHA--VSGVFATVADAVITPMDVVKQRLQSS 163  
OY 187 SYRELGACVTRFVAVAGGKRSIMLGWPTALRDVPSALYWFYELVKSMNLGRPK--DQ 244  
DB 164 PYKGVVDCVRRLVEEGGAFYASRTTYVMNAPPTAVHFAFYETAKKGLLEVSEPTAND 223  
OY 245 TSVGNSFVAGISGTVAALVLPEDVVKTORQVALGAEAVRNPLHYDSTWLLLRIRRA 304  
DB 224 ENLVHATAGAGAAALAAVVTPLDVVKTLQ-CQVCGCDRFSSTIOD--VIGSIYK 279  
OY 305 ESGTGLFAGFLPRTIKAPSCAISMISTYERKSFQRLNQ 345  
DB 280 KNGYGLMRGWIPRLFHAPAALICWSTYEASKTFEOKLINE 320

RESULT 4  
US-09-068-140A-10  
Sequence 10, Application US/09068140A  
Patent No. 6281409

GENERAL INFORMATION:  
APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor  
APPLICANT: and Rex Michael Brennan  
TITLE OF INVENTION: Blackcurrant Promoters and Genes  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/068.140A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP96/04807  
FILING DATE: No. 6281409ember 4, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Dimer, Dara L.  
REGISTRATION NUMBER: 33,680  
REFERENCE/DOCKET NUMBER: C70237  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5017  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 289 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Rides nigrum  
STRAIN: Ben Alder  
US-09-068-140A-10

Query Match 17.5%; Score 320; DB 4; Length 289;  
Best Local Similarity 26.9%; Pred. No. 7.5e-25;  
Matches 90; Conservative 64; Mismatches 124; Indels 56; Gaps 7;

OY 15 MVASGTGAVVTSLEFPTPLDVVKVRLQSORPSMASLMSRLMSLYTKMKCLLYCNVVL 74  
DB 1 MIAGSIAGSIEHMAHYVDLTKRIQ-----IGSCS----- 32  
OY 75 EPLYLCNGARCAWFOPTREFTGMDAFKIVRHEGTRTLMGSLPATLVMTVATYF 134  
DB 33 -----AQSAGLRQALGSLIKKEGPAGLYRGIGAMGLGAPHAHYF 73  
OY 135 TAYDQK-AFLCGRALTSDLAPMYAGALRLGTYTVISPLEMRTKIQAOHVSYRELGA 193  
DB 74 SYELCKEFTFSGHGPSNSGAHA--VSGVFATVADAVITPMDVVKQRLQSSPYKGVVD 131  
OY 194 CVRTVAVAGGKRSIMLGWPTALRDVPSALYWFYELVKSMNLGRPK--DQTSVGSF 251  
DB 132 CVRRLVEEGGAFYASRTTYVMNAPPTAVHFAFYETAKKGLLEVSEPTANDLVHA 191  
OY 252 VAGISGTVAALVLPEDVVKTORQVALGAEAVRNPLHYDSTWLLLRIRAESGTGKL 311  
DB 192 TAGAAGALAAVVTPLDVVKTLQ-CQVCGCDRFSSTIOD--VIGSIYKKNKYGL 247  
OY 312 FAGFLPRTIKAPSCAISMISTYERKSFQRLNQ 345  
DB 248 MRGWIPRLFHAPAALICWSTYEASKTFEOKLINE 281

RESULT 5  
US-08-518-878B-56  
Sequence 56, Application US/08518878B  
Patent No. 5702802

GENERAL INFORMATION:  
APPLICANT: Tartaglia, Louis A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York

COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/518,878B  
FILING DATE: 23-AUG-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-036  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9090  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 299 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
US-08-518-878B-56

Query Match 13.0%; Score 237; DB 1; Length 299;  
Best Local Similarity 25.1%; Pred. No. 2.7e-16;  
Matches 85; Conservative 43; Mismatches 132; Indels 78; Gaps 10;

QY 14 QWASGTGAVVTSLEMTPLDVVKVRLQSORPSMASLSPSSRLMSLSYTKMKCLLYCNGV 73  
DB 6 KRLGAGTACIADLTPLDFAKVRQIGES----- 37  
QY 74 LEPVLYCPNGACATWFOPTPTFT-----GTMDAFVKIVRHGRTILMSGIPALIVMTV 127  
DB 38 -----QGPVATVSAOYRGVGTILMVREGPRSLYNGLVAGLQROM 80  
QY 128 PATAIYFAYDOLKAFLC-----GRAITSDLYAPMVAGALARLGTWVVISPLEMRKLOA 183  
DB 81 SFAVSRIGLYDSVKQFYTKGSEHASISGRLLAGSTTGALA-----VAVAOPTVYVVRQA 136  
QY 184 QHVS-----YRELGACVRTVAOAGWRSIMLGMGPTALRDVPSALYMWENYELVSWLNG 238  
DB 137 QARAGGRRYQSTVNAVKTIAREEGFRGLMKGTSNVANALVNCALVYDILDAL-- 194  
QY 239 LRPKQOT-SVGSFVAAGISGTVAAVLTPEDVYKTO-QOVALGAMEAVRNPPLHVDSTW 296  
DB 195 LKANIMTDDLPCHTFSAFGAGCTTVIASPVVDVVKTRYNSALGOYSSAG---HCAITM 250  
QY 297 LLRLRIAESGTGKGFAGFLPRIIRAKAPSCAIMISTYE 334  
DB 251 L-----QKSGPRATYKGFMPFSLRLGSMNVVMFYTYE 282

RESULT 6  
US-08-470-868A-56  
Sequence 56, Application US/08470868A  
Patent No. 5861485  
GENERAL INFORMATION:  
APPLICANT: Tartaglia, Louis C.  
TITLE OF INVENTION: Compositions and Methods for the  
TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie and Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/470,868A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7853-0031-999  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864  
TELEX: 66441 PENNIE  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 299 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
US-08-470-868A-56

Query Match 13.0%; Score 237; DB 2; Length 299;  
Best Local Similarity 25.1%; Pred. No. 2.7e-16;  
Matches 85; Conservative 43; Mismatches 132; Indels 78; Gaps 10;

QY 14 QWASGTGAVVTSLEMTPLDVVKVRLQSORPSMASLSPSSRLMSLSYTKMKCLLYCNGV 73  
DB 6 KRLGAGTACIADLTPLDFAKVRQIGES----- 37  
QY 74 LEPVLYCPNGACATWFOPTPTFT-----GTMDAFVKIVRHGRTILMSGIPALIVMTV 127  
DB 38 -----QGPVATVSAOYRGVGTILMVREGPRSLYNGLVAGLQROM 80  
QY 128 PATAIYFAYDOLKAFLC-----GRAITSDLYAPMVAGALARLGTWVVISPLEMRKLOA 183  
DB 81 SFAVSRIGLYDSVKQFYTKGSEHASISGRLLAGSTTGALA-----VAVAOPTVYVVRQA 136  
QY 184 QHVS-----YRELGACVRTVAOAGWRSIMLGMGPTALRDVPSALYMWENYELVSWLNG 238  
DB 137 QARAGGRRYQSTVNAVKTIAREEGFRGLMKGTSNVANALVNCALVYDILDAL-- 194  
QY 239 LRPKQOT-SVGSFVAAGISGTVAAVLTPEDVYKTO-QOVALGAMEAVRNPPLHVDSTW 296  
DB 195 LKANIMTDDLPCHTFSAFGAGCTTVIASPVVDVVKTRYNSALGOYSSAG---HCAITM 250  
QY 297 LLRLRIAESGTGKGFAGFLPRIIRAKAPSCAIMISTYE 334  
DB 251 L-----QKSGPRATYKGFMPFSLRLGSMNVVMFYTYE 282

RESULT 7  
US-08-518-878B-51  
Sequence 51, Application US/08518878B  
Patent No. 5702902  
GENERAL INFORMATION:  
APPLICANT: Tartaglia, Louis A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 9  
 US-08-670-868A-51  
 ; Sequence 51, Application US/08470868A  
 ; Patent No. 5861485  
 ;  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tartaglia, Louis C.  
 ; TITLE OF INVENTION: Compositions and Methods for the  
 ; TITLE OF INVENTION: Treatment of Body Weight Disorders, Including Obesity  
 ; NUMBER OF SEQUENCES: 56  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie and Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA

ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/470,868A  
 FILING DATE: 06-JUN-1995  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Coruzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7853-0031-999  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 790-9090  
 TELEFAX: (212) 869-8864  
 TELEX: 66441 PENNIE  
 INFORMATION FOR SEQ ID NO: 51:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 309 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: unknown  
 US-08-470-868A-51

Query Match 13.0%; Score 237; DB 2; Length 309;  
 Best Local Similarity 25.1%; Pred. No. 2.8e-16;  
 Matches 85; Conservative 43; Mismatches 132; Indels 78; Gaps 10;

QY 14 QMVASGTGAVVTSLEMTPLDVYKVRLOSQPSMASSELMSSRLMSLSYTKMKCLLYCNGV 73  
 DB 16 KFLGAGTACTADLTLPDLTAQVRLQIGES----- 47  
 QY 74 LEPLVLCPNRGARCATWFOPTPTFT-----GTMDAFVKYVHEGRTTMSGLPATLVMTV 127  
 DB 48 -----GQVRAVTSAGYRGVMTGLTMTWTEGPRSLXNGVAGLQROM 90  
 QY 128 PATATYTTAVDOLKAFLC-----GRALTSGLYAPMVAGALARIAGTVVVISPLELMTKLOA 183  
 DB 91 SFASVRLGLDVKQFPTKSGEHSASISRLLAGSTGALA-----VAAQPTDVYKVRFOA 146  
 QY 184 QHVS-----YRELGACVRTAVAAQGWRSMLTGMGPTALRDVPFSALYWFNYELVKSMLNG 238  
 DB 147 QARAGGRRTQSTYNAVKTIAREEGFRGLMKGTSPNVARNAIVNCAELVYDILKDAL-- 204  
 QY 239 LRPKDOT-SYGMSTVAGGISTYAAVLTLPDVYKTO-ROYALGAMEAVRNPPLHVDSTW 296  
 DB 205 LKANIMTDDLPCHTSAFGAGFCTTVIASPVDVYKTRYNMSALGOYSAG-----HCALTM 260  
 QY 297 LLRRIRAESGTGKLFAGFLPRIRIKAAPSCAISMISTYE 334  
 DB 261 L-----QKEGPRAFYKGFMPFLRLGSMNVVMFVTE 292

RESULT 10  
 US-09-210-681-51  
 Sequence 51, Application US/09210681  
 Patent No. 6057109  
 GENERAL INFORMATION:  
 APPLICANT: Tartaglia, Louis A.  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE  
 REGULATION OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
 NUMBER OF SEQUENCES: 64  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds LLP  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036-2711  
 COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/210,681  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/807,861  
 FILING DATE: 26-FEB-1997  
 APPLICATION NUMBER: US 08/518,878  
 FILING DATE: 23-AUG-1995  
 APPLICATION NUMBER: US 08/470,868  
 FILING DATE: 06-JUN-1995  
 APPLICATION NUMBER: US 08/294,522  
 FILING DATE: 23-AUG-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Coruzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7853-066  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 790-9090  
 TELEFAX: (212) 869-9741/8864  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 51:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 309 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: unknown  
 US-09-210-681-51

Query Match 13.0%; Score 237; DB 3; Length 309;  
 Best Local Similarity 25.1%; Pred. No. 2.8e-16;  
 Matches 85; Conservative 43; Mismatches 132; Indels 78; Gaps 10;

QY 14 QMVASGTGAVVTSLEMTPLDVYKVRLOSQPSMASSELMSSRLMSLSYTKMKCLLYCNGV 73  
 DB 16 KFLGAGTACTADLTLPDLTAQVRLQIGES----- 47  
 QY 74 LEPLVLCPNRGARCATWFOPTPTFT-----GTMDAFVKYVHEGRTTMSGLPATLVMTV 127  
 DB 48 -----GQVRAVTSAGYRGVMTGLTMTWTEGPRSLXNGVAGLQROM 90  
 QY 128 PATATYTTAVDOLKAFLC-----GRALTSGLYAPMVAGALARIAGTVVVISPLELMTKLOA 183  
 DB 91 SFASVRLGLDVKQFPTKSGEHSASISRLLAGSTGALA-----VAAQPTDVYKVRFOA 146  
 QY 184 QHVS-----YRELGACVRTAVAAQGWRSMLTGMGPTALRDVPFSALYWFNYELVKSMLNG 238  
 DB 147 QARAGGRRTQSTYNAVKTIAREEGFRGLMKGTSPNVARNAIVNCAELVYDILKDAL-- 204  
 QY 239 LRPKDOT-SYGMSTVAGGISTYAAVLTLPDVYKTO-ROYALGAMEAVRNPPLHVDSTW 296  
 DB 205 LKANIMTDDLPCHTSAFGAGFCTTVIASPVDVYKTRYNMSALGOYSAG-----HCALTM 260  
 QY 297 LLRRIRAESGTGKLFAGFLPRIRIKAAPSCAISMISTYE 334  
 DB 261 L-----QKEGPRAFYKGFMPFLRLGSMNVVMFVTE 292

RESULT 11  
 US-08-946-719A-51  
 Sequence 51, Application US/08946719A  
 Patent No. 6121017  
 GENERAL INFORMATION:  
 APPLICANT: Tartaglia, Louis A.  
 TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
 REGULATION OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
 NUMBER OF SEQUENCES: 64  
 CORRESPONDENCE ADDRESS:

```

: ADDRESSEE: Pennie & Edmonds LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/946,719A
: FILING DATE: 8-OCT-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/807,861
: FILING DATE: 26-FEB-1997
: APPLICATION NUMBER: US 08/518,878
: FILING DATE: 23-AUG-1995
: APPLICATION NUMBER: US 08/470,868
: FILING DATE: 06-JUN-1995
: APPLICATION NUMBER: US 08/294,522
: FILING DATE: 23-AUG-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7853-107
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEFAX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 51:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 309 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: US-08-946-719A-51

```

```

Query Match 13.0%; Score 237; DB 3; Length 309;
Best Local Similarity 25.1%; Pred. No. 2.8e-16;
Matches 85; Conservative 43; Mismatches 132; Indels 78; Gaps 10;

QY 14 OMVAGSTGAVVNTSLFMTPLDVVKVRLQSORPMSASELMPSSRLMSLSTKKMKCLYKGV 73
DB 16 KFLGAGTACIADLTITPLDTAKVKRLQIGES----- 47
QY 74 LEPLVLCPRGACATWFOPTREF-----GTMDAFVKIVRHEGRTLMGSLPATLVMTV 127
DB 48 -----GGPVRAIVSAQYRGVAGTILTFVTRPREGRSIYNGVLQGRQM 90
QY 128 PATATATTAAYDOLKAFIC-----GRALTSPLVAPWAGALARIQIVYISPLELMTKQIA 183
DB 91 SPASVRIGLYDVVKFYTKGSEHASISGRLLAGSTTGALA-----VAVAPPTVAVVRQIA 146
QY 184 OHVS-----YRELGACVTRAVAGGWRSLMLGMPALRDVPEFSALVWYNTYELVSWLNG 238
DB 147 QARAGGRRYQSTVNVAIKTIAREEGFRGLMKSTSPRVARNALVNCALVYTLIDAL-- 204
QY 239 LRPKDOT-SVGMSEVAGGISGTVAAVLTLPLFDVVKTO--QOVALGMEAVRNPPLVDSTW 296
DB 205 LKANLMTDOLPCHFTSAFGAGCTTVIASPVVDVVKTRYNNSALGQYSAG-----HCALTM 260
QY 297 LLRLRIASSTGKGLFAGLPLRIIKAAPSCAIMISTYE 334
DB 261 L-----QKEGPRAFYKGMPSFLRLGSLWNVVMEVYIE 292

```

RESULT 12  
US-08-933-750C-12  
Sequence 12, Application US/08933750C

```

: Patent No. 5932442
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Bandman, Olga
: APPLICANT: Shah, Puryl
: APPLICANT: Au-Young, Janice
: APPLICANT: Yue, Henry
: APPLICANT: Guegler, Karl J.
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
: NUMBER OF SEQUENCES: 98
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/933,750C
: FILING DATE: September 23, 1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0356 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEFAX:
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 320 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: SPINNOT02
: CLONE: 207452
: US-08-933-750C-12

```

```

Query Match 12.9%; Score 236; DB 2; Length 320;
Best Local Similarity 22.0%; Pred. No. 3.7e-16;
Matches 83; Conservative 58; Mismatches 136; Indels 100; Gaps 12;

QY 3 DDDPAG--ISPLQOMVAGSTGAVVNTSLFMTPLDVVKVRLQSORPMSASELMPSSRLMSLS 60
DB 5 DKPDRNNNTKQVAVASVSGIVTRALISPDVYIKIRFOLEHRLSR----- 52
QY 61 YTKMKCLYKCNVLEPLVLCPRGACATWFOPT--RFTGMDAFVKIVRHEGRTLMGSL 119
DB 53 -----SDPSAKYHGIQASQIIOEBEPTLFMKGH 82
QY 120 PATLVTPATATTAAYDOLKA-----FLGRALTSPLVAPWAGALA 163
DB 83 VPAQILSTIGYAVQPLSEMLELTVHRSYVDAREFSVHPFG-----GLAA 129
QY 164 RLGVTVYISPLELMTKQIAOHVS--YRELGACVTRAVAGGWRSLMLGMPALRDVVPF 221
DB 130 CWATLTV-HPVDVLTTRPFAOGEPRVYNTLIRAVGTWYRSEGPQVFGKGLAPTLTAIRPY 188
QY 222 SALVWFNVELVK-----SWLNGLRPKDOTSVGMSEVAGGISGTVAAVLTLPLFDVVKTOROVA 278

```

Db 189 AGIOFSCYSLSKHLKWLPAESGKNENL-ONLLCGSGAGVISKTLTYPLDLFKRLQV- 246  
 QY 279 LGAMEAVRNPLHVDSTWLLLRIRA-----ESGTKGLFAGFLPRITKAAPSCAI 328  
 Db 247 -GGFE-----HARAFQGVRRYKGLMDCAKOVLOKEGALGFYKLSPLSKAALSTGE 298  
 QY 329 MISTYEFKSFQRLNQ 345  
 Db 299 MFSYEFQCNVPHCMNR 315

## RESULT 13

US-09-234-613-12  
 ; Sequence 12, Application US/09234613  
 ; Patent No. 6132973  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lal, Preeti  
 ; APPLICANT: Hillman, Jennifer L.  
 ; APPLICANT: Bandman, Olga  
 ; APPLICANT: Shah, Purvi  
 ; APPLICANT: Au-Young, Janice  
 ; APPLICANT: Yue, Henry  
 ; APPLICANT: Guegler, Karl J.  
 ; APPLICANT: Corley, Neil C.  
 ; TITLE OF INVENTION: HORMAN REGULATORY MOLECULES  
 ; NUMBER OF SEQUENCES: 98  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/234,613  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/933,750  
 ; FILING DATE: September 23, 1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 12:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 320 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: SPINNOT02  
 ; CLONE: 207452  
 ; US-09-234-613-12

Query Match 12.9%; Score 236; DB 4; Length 320;  
 Best Local Similarity 22.0%; Pred. No. 3.7e-16;

Matches 83; Conservative 58; Mismatches 136; Indels 100; Gaps 12;

QY 3 DDPAG--ISPLQMVASGTAVVTSLEMTPLDVVKVRLQSGRPSMASLMPSSRLMSLS 60  
 Db 5 DPKPDGRNNTKFOYAVVAGSVSGLVTRALISPPDVIKIRFOQHEHLSR----- 52

QY 61 YKKMKCLLYCNGVLEBPLYLCPNGARCATWFOPT-RTGTMDAFVKYVHEGTRTMSGL 119  
 Db 53 -----SDBSAKYHGLQASRQLQEGPTAWKGH 82  
 QY 120 PATLVNTVPATAIYTPAQDLKA-----FLGRALTSPLVAPVAGALA 163  
 Db 83 VPAQIISIGYGAQVFLSFEMLTETELHRGSVYDAREFSVHFVCG-----GLVA 129  
 QY 164 RLGTVTVPISPLEMTKRLQAOHVS--YRELGACVTRAVAOGGMRSLMLGPTALRDVPE 221  
 Db 130 CMATLTIV-HPVDLTKTRFPAAGPEPKYNTLRHNAVGTMYRSEGPVYKGLATPTLIAIFPY 188  
 QY 222 SALWMENYELVK---SWLNGLRPKDOTSVGMSFVAGISGTVAAYVLTLPFDVYKTQROVA 278  
 Db 189 AGIOFSCYSLSKHLKWLPAESGKNENL-ONLLCGSGAGVISKTLTYPLDLFKRLQV- 246  
 QY 279 LGAMEAVRNPLHVDSTWLLLRIRA-----ESGTKGLFAGFLPRITKAAPSCAI 328  
 Db 247 -GGFE-----HARAFQGVRRYKGLMDCAKOVLOKEGALGFYKLSPLSKAALSTGE 298  
 QY 329 MISTYEFKSFQRLNQ 345  
 Db 299 MFSYEFQCNVPHCMNR 315

## RESULT 14

US-09-188-930-339  
 ; Sequence 339, Application US/09188930A  
 ; Patent No. 6150502  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Watson, James D.  
 ; APPLICANT: Strachan, Lorna  
 ; APPLICANT: Sleeman, Matthew  
 ; APPLICANT: Onrust, Rene  
 ; APPLICANT: Murison, James Greg  
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
 ; FILE REFERENCE: 11000.1011c1  
 ; CURRENT APPLICATION NUMBER: US/09/188,930A  
 ; CURRENT FILING DATE: 1998-11-09  
 ; NUMBER OF SEQ ID NOS: 348  
 ; SOFTWARE: FASTSEQ for Windows Version 3.0  
 ; SEQ ID NO 339  
 ; LENGTH: 469  
 ; TYPE: PRT  
 ; ORGANISM: Mouse  
 ; US-09-188-930-339

Query Match 12.8%; Score 233.5; DB 4; Length 469;

Best Local Similarity 23.2%; Pred. No. 1.2e-15;

Matches 77; Conservative 59; Mismatches 125; Indels 71; Gaps 9;

QY 13 QOMVASGTAVVTSLEMTPLDVVKVRLQSGRPSMASLMPSSRLMSLSYTKKCLLYONG 72  
 Db 188 RHLVAGGAGAVSRTCTAPLRLKVLMO-----VHASRSNNM-----CI----- 226  
 QY 73 VLEPLYLCPNGARCATWFOPTRTGTMDAFVKYVHEGTRTMSGLPATLVNTVPATAI 132  
 Db 227 -----VGGFTQMIREGCAKSLMRGNKINVLKAPESAI 259  
 QY 133 YFTAYDQKAPLQGRALTSLSLYAPVAGALARLGTIVISPLEMTKRLQAOHV-SYREL 191  
 Db 260 KEMAVEQMKRLVGSPOETLRIHERLVAGSLAGALAOSSIYPMVEVLKTRMALRKQGYSGM 319  
 QY 192 GACVTVTAOOGMSLWGMGPTALRDVPSALWTFNELYK-SWL-----NGLRPKDOT 245  
 Db 320 LDCARRILAKGVAAFYKGIIPNMGIIPYAGIDLAVETLKNWLOXYAVNASDP----- 375  
 QY 246 SVGMSFVAGISGTVAAYVLTLPFDVYKTQROVAL---GAMDAVRNPLHVDSTWLLLR 302  
 Db 376 GVFTVLACGTISTTCGOLASTPLALVTRMQAQSIBGAPVYMS-----LKKQI 426

OY 303 RAESGTGGLFAGFLPRITKAPSCAIMISTYE 334  
Db 427 LRTEGAGLRYGLAPNFKVTPAVSISIVYE 458

## RESULT 15

US-08-937-466-2  
Sequence 2, Application US/08937466  
Patent No. 5846779  
GENERAL INFORMATION:  
APPLICANT: Zhang, Ning  
APPLICANT: Amaral, M. Catherine  
TITLE OF INVENTION: UCP3 Genes  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 75 DENISE DRIVE  
CITY: HILLSBOROUGH  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94010  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/937,466  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: T97-009  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 343-4341  
TELEFAX: (650) 343-4342  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 308 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-937-466-2

Query Match 12.1%; Score 221.5; DB 2; Length 308;  
Best Local Similarity 24.0%; Pred. No. 1.1e-14;  
Matches 83; Conservative 49; Mismatches 147; Indels 67; Gaps 9;

OY 1 MADDOPAGISP--LQOMVASTGAVVTSLEMTPLDVKVRLQOSQPSMASSEIMPSRLMS 58  
Db 1 MYGLQPSSEVPPTVVKFAGAGTACFAADLTFPLDTAKVRLQIOGPNQAQ----- 51  
OY 59 LSTKMKCLLYCNGVLEPLVLCPNBARCATWFOPTREFTGMDAFVKIVRHGSTRITMSG 118  
Db 52 -----SVQTRGVLGITLTWVTRTEGPRSPYSG 77  
OY 119 LPATLVMTVPATAYFYAVDOLKAFLCGR-ALTSDLXAPVAVGALARLGTVVISPLEIM 177  
Db 78 LVAGLHRQMSFASIRIGLYDSVKOFTPKGADHSSVAIRILAGCTGAMAVTCAQPTDV 137  
OY 178 RTKLQAO-----HVSRELGACVFTAVVAVGGRSLMLGMPALADVPFSALYWFNYE 230  
Db 138 KVRFOAMIRLCTGGERKTRGTMDAYRTIAREGVRLMKGTWPNITRNAIVNCAEWYTD 197  
OY 231 LVKSWLNGLRPKDQT-SVGSSEVAVAGISGTVAVALTLPFDVVKTOROVA-LGAEAVRVN 288  
Db 198 IIKKEL--LBSHLFTDNFPCHFVSAFGACATVVASPDVVKRYNNAPLGRYR-----S 251  
OY 289 PLHVDSTWLLRLRRASGTGKGLFAGFLPRITKAPSCAIMISTYE 334

Db 252 PLH-----CMKMAAQEGPTAFYKGFVPSFLRLGAMVNMMEVTE 291

Search completed: April 19, 2002, 22:41:59  
Job time: 3357 sec

• • • •

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2002, 20:34:46 ; Search time 172.76 Seconds  
(without alignments)  
8153.423 Million cell updates/sec

Title: US-09-840-787-68  
Perfect score: 1643  
Sequence: 1 GCGAGCGCTGACGACGAGAA.....GATGACTTCACAAAAA 1643

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues  
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N\_Geneseq\_1101.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	1449	88.2	1560	21	AAC90452	Human uncoupling P	
2	1405.4	85.5	1662	21	AAC76992	Human ORF2547	
3	532	32.4	537	21	AAA06606	Human Immunogenic	
4	532	32.4	537	22	AAH93722	Human prostate-spe	
5	532	32.4	537	22	AAH85036	Human prostate-spe	
6	532	32.4	537	22	AAH02787	Prostate tumour an	
c	7	243.8	14.8	6712	21	AAC76691	Human ORF2246
8	168.4	10.2	174	16	AAT26018	Human gene signatu	
9	144.2	8.8	397	21	AAH30252	Human colon cancer	
c	10	74.8	4.6	452	22	AA139053	Probe #7739 used t
11	74.8	4.6	496	22	AA139123	Probe #7809 used t	

c	12	65.4	4.0	72604	20 AA210752	Genomic sequence o
c	13	64.2	3.9	4231	21 AAC77105	Human ORFX ORF2660
c	14	63.6	3.9	444	22 AAF65290	Novel human polyu
c	15	63.4	3.9	545	22 AAH09622	Human cDNA clone (
c	16	63.4	3.9	1517	22 AAH17208	Human cDNA sequenc
c	17	63	3.8	5216	22 AAH28355	Nucleotide sequenc
c	18	62.8	3.8	4089	20 AA208807	Human cerebral pro
c	19	62.8	3.8	87350	18 AA83003	Human WRN genomic
c	20	62	3.8	5009	19 AA65634	First EcoRI nucleo
c	21	62	3.8	26928	20 AA232184	Human prothrombin
c	22	62	3.8	48000	22 AAF27996	Human calcium sens
c	23	61.8	3.8	2492	21 AAC78199	Human cancer assoc
c	24	61.6	3.7	700	22 AAH92306	Human inflammatory
c	25	61.6	3.7	1013	21 AAC68083	Human secreted pro
c	26	61.6	3.7	7680	21 AA89439	14-3-3 sigma trans
c	27	61.6	3.7	9439	22 AAF3723	Human kidney relat
c	28	61.6	3.7	50000	21 AA86363	Polymorphic repeat
c	29	61.6	3.7	97662	22 AAF83908	Genomic sequence o
c	30	61.6	3.7	106746	21 AA810225	Human PCNA-1 genom
c	31	61.4	3.7	215	21 AAC29167	Human secreted pro
c	32	61.4	3.7	1545	21 AAC59037	Human secreted pro
c	33	61.4	3.7	7680	21 AA89439	14-3-3 sigma trans
c	34	61.2	3.7	2202	22 AAH17401	Human cDNA sequenc
c	35	61.2	3.7	17858	22 AA164114	Human bladder rela
c	36	61.2	3.7	160552	22 AAD02697	Human glycosyl sul
c	37	61	3.7	323	21 AAC07236	Human secreted pro
c	38	61	3.7	997	21 AA261628	cDNA encoding murI
c	39	61	3.7	997	22 AAC99561	SKIN cell cDNA, SE
c	40	61	3.7	1808	22 AAC99806	SKIN cell cDNA, SE
c	41	61	3.7	1816	21 AA261789	cDNA encoding murI
c	42	61	3.7	1816	22 AAC99722	SKIN cell cDNA, SE
c	43	61	3.7	2270	22 AAH16683	Human cDNA sequenc
c	44	61	3.7	110000	22 AAF84800	Nucleotide sequenc
c	45	61	3.7	161425	22 AAH02340	Human ARAPI0 gene

ALIGNMENTS

RESULT 1	
AAC90452	standard; cDNA: 1560 BP.
AC	AAC90452;
XX	
DT	12-MAR-2001 (first entry)
XX	
DE	Human uncoupling protein cDNA #1.
XX	
KW	Human; uncoupling protein; immunosuppressive; antiarthritic;
KW	antiinflammatory; antiproliferative; cardiant; vasotropic;
KW	cerebroprotective; neuroprotective, antibacterial; ophthalmological;
KW	gastrointestinal; nephrotropic; gynecological; vulvneary; thrombotic;
KW	gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;
XX	Intertility; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200061614-A2.
XX	
PD	19-OCT-2000.
XX	
PF	06-APR-2000; 2000WO-US09534.
XX	
PR	09-APR-1999; 99US-0128701.
PR	08-JUL-1999; 98US-0142821.
PR	18-AUG-1999; 99US-0149448.
XX	12-NOV-1999; 99US-0164751.
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;
XX	

DR WPI: 2000-656322/63.  
 DR P-PSDB: AAB50378.  
 XX Uncoupling proteins and nucleic acid sequences encoding them, useful  
 PT for detecting, preventing and treating proliferative, neurological,  
 PT immune system, cardiovascular and gastrointestinal disorders -  
 XX  
 PS Claim 1; Page 303; 343pp; English.  
 XX  
 CC The present sequence is one of eighteen isolated nucleotide sequences  
 CC encoding uncoupling proteins. The nucleotide sequences may be used for  
 CC the detection of various disorders such as cancer, for chromosome  
 CC identification, as chromosome markers and for numerous other diagnostic  
 CC or research purposes. The uncoupling protein encoded by the nucleotide  
 CC sequences may be used to treat disorders such as neural, immune,  
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
 CC renal and proliferative disorders, wounds, infectious diseases,  
 CC thrombosis, arthritis, and infertility.  
 CC  
 SO Sequence 1560 BP; 286 A; 492 C; 459 G; 323 T; 0 other;

Query Match 88.2%; Score 1449; DB 21; Length 1560;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1452; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 187 CAGAGCTAAGCTTCAAGATGGCTGACGAGACCTGGGGGATGAGCCCTCCAGCAA 246  
 DB 102 cagagctgaagcttcaagatggtgaccagacccctgagatcagccccctccagcaa 161  
 QY 247 ATGGTGGCTCAGGACCGGGGCTGTGTTACTCTCTCTTATACACACCTTGACGTG 306  
 DB 162 atggtggctcagagacccgggctgtgttactctcttcaacccccctgagctg 221  
 QY 307 GTGAGTGTGCGCTCAGTCTCAGCGGCGCTTCATGGCGAGAGCTGATGCTTCCTCC 366  
 DB 222 gtgaaagtgtcgtcgtcagatctcagcgcctccatgagcagagagctgagcttcc 281  
 QY 367 AGACTGTGAGAGCTCTCTTATACCAATGGAAGTGCCTCTGTATTTGCAATGGTCTCG 426  
 DB 282 agactgtgagagctctcttatacaaatggaagtgcctctgtatctgcaatgtgtcctg 341  
 QY 427 GAGCTCTGTACTGTGCGCAATGATGGCTGTGCGCTGTGCGACCTGGTTTCAGACCTTAC 486  
 DB 342 gagctctgtactgtgctcgaatgtgtccgctgtgcaacctgttcaagacctacc 401  
 QY 487 CGCTCAGTGGAGCAGATGATGCTTCGTGAAGATCGTAGAGCAGAGGACACGAGACC 546  
 DB 402 cgcttcactgacacatgtagtgcctgtgaagatcgtgagcagagagggccagagacc 461  
 QY 547 CTGTGAGCGGCGCTCCCGCACCTGTGTGATGACTGTGCGAGTACCGCATTTACTTC 606  
 DB 462 ctgtgagcggtccctccgcacacctgtgtgactgtgcaagctaccgcaactacttc 521  
 QY 607 ACTGCTTATGACCAACTGAAGGCTTCTGTGTGTGTCAGGCCCTTACTCTGACTCTAC 666  
 DB 522 actgcttatgaccaactgaaggcttctgtgtgtgtcagaccctgaacctgaacctacc 581  
 QY 667 GCACCATGATGGCTGGGCGCTGGCCGCTGGGACCGTAGATGTATAGACCCCGCTG 726  
 DB 582 gcaacctatggtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 641  
 QY 727 GAGCTTATCGGACAAAGCTGACAGCTGACATGTGTACCGGAGAGCTGGTCTGT 786  
 DB 642 gagcttatcgacaagaactgcaagctcagcatgtgtctaccggagctgtgtggtggt 701  
 QY 787 GTTGAATGCAATGCTCAGAGGTGCTGCGCTACTGTGCTGTGGCTGGGGCCCACT 846  
 DB 702 gttaactgcaatgctcagaggtgctgctgctgctgctgctgctgctgctgctgct 761  
 QY 847 GCCCTGAGATGAGCTTCTGACAGCCCTGTACTGCTCACTAAGACTGAGCTGTAAGAGC 906  
 DB 762 gcccttgagatgagcttctcagccctgtactggttcaactatgagctgtgtgagaagc 821

QY 907 TGGCTCAATGGGCTCAGCGGAGGACGAGCTTGTGTGGGATGAGCTTTGTGCTGT 966  
 DB 822 tggctcaatgggctcagcgaggagcagagcagacttctgtgtgcatgagcttgtgtgct 881  
 QY 967 GGCATCTCAGGAGAGGCTGGCTGACGTGACTCTACCTTTGAGCTGTGTAAGACCCA 1026  
 DB 882 ggcattctcagagagagctggctgacgtgacttacccttgaagctgtgtaagacc 941  
 QY 1027 CGCAGGTGCTCTGTGGAGCAGTGAAGCTGTGAGAGTGAACCCCTGATGTGACTTC 1086  
 DB 942 cgcaggtgctctgtggagcagatgagctgtgagagtgaaacccctgcatgtgacttc 1001  
 QY 1087 ACCGTGCTCTCTCGGAGAGATCCGGGCGAGTGGGCGACCAAGGACTTTTGCAAGC 1146  
 DB 1002 accgtgctctctcggagagatccgggcgagtggcgagcccaagagacttctgagc 1061  
 QY 1147 TTCCTCCCGGATCATCAAGGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1206  
 DB 1062 ttcctcccgatcatcaaggctgccctgctgctgctgctgctgctgctgctgctgct 1121  
 QY 1207 TTCGCAAAAGCTTCTTCAGAGAGCTGAACAGACCGGCTTCTGGGCGCTGAAGGG 1266  
 DB 1122 ttcgcaaaagcttcttcagagagctgaacacggagcagcgttctgtggtcgctgaagg 1181  
 QY 1267 CAGAGAGGCAAGACCCGCTCTCTCCACGATGGGAGAGGGCAAGAGACCCAGCC 1326  
 DB 1182 caagagagcaagacccgctctctccacgatlgygagagagcagagagaccagcc 1241  
 QY 1327 AAGTGCCTTTCTTCAGACATGAGGAGGAGGCTGTGCTTCTCCCTCCGCGGACAG 1386  
 DB 1242 aagtgccttcttcagacatgagagagagagagctgttccctccctcccgagcaag 1301  
 QY 1387 CTCAGAGGAGGAGGCTGCTCCTGTGGGCGGCGCAGACTTCTCAGACAACTTCTCT 1446  
 DB 1302 ctcaagagagagagctgctcctctgtggcgccagacttccctccagacaacttctct 1361  
 QY 1447 GCTGCTCAGTGTGGGATCATCACTTACCCACCCCAAGTTGAGACCAATCTTCT 1506  
 DB 1362 gctgctcagtggtgagatcaactaccaccccaagctcaagaaccaactctcc 1421  
 QY 1507 AGCTGCCCCCTGCTGCTTCCCTGTGTTGCTGTAGCTGGGATGCTTCACAGACCA 1566  
 DB 1422 agctgccccctgctgcttccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1481  
 QY 1567 AAGCCCTAGCTGTGATGCTCCCTGACCTTTAATCTTAAGTCTAAGATGAT 1626  
 DB 1482 aagccctagctgtgagatctccctgaacctgttaattctcaagtcaagaatgat 1541  
 QY 1627 GAATTCAAAAAAA 1643  
 DB 1542 gaaaaaataaaaaa 1558

RESULT 2  
 AAC76992  
 ID AAC76992 standard; cDNA; 1662 BP.  
 XX  
 AC AAC76992;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORF5247 polynucleotide sequence SEQ ID NO:5093.  
 XX  
 KW Human; open reading frame; ORF; detection; cytosolic; hepatotropic;  
 KW vlnular; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiast;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;  
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;

KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive; ss.  
 XX Homo sapiens.  
 OS  
 PN MO200058473-A2.  
 PD 05-OCT-2000.  
 XX  
 XX 31-MAR-2000; 2000WO-US08621.  
 PF  
 XX 31-MAR-1999; 99US-0127607.  
 PR 02-APR-1999; 99US-0127636.  
 PR 05-APR-1999; 99US-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shmukets RA, Leach M;  
 DR WPI: 2000-602362/57.  
 DR P-PSDB: AAB42783.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 XX  
 XX Claim 5: Page 4281-4282; 5507pp; English.  
 XX  
 XX AACT74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORF open reading frames 1 to 3161. The ORF  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;  
 CC antiproliferic; antiparkinsonian; nootropic; immunosuppressive;  
 CC osteopathic; anticonvulsant; antiallergic; immunosuppressive;  
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;  
 CC antihypoid; and antianemic. The sequences can be used for determining  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORF-associated disorder. The  
 CC nucleic acids can be used to express ORF proteins in gene therapy  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
 CC coagulation; to inhibit thrombosis; and as a contraceptive.  
 XX  
 XX Sequence 1662 BP; 336 A; 516 C; 468 G; 341 T; 1 other;  
 SQ

Query Match 85.5%; Score 1405.4; DB 21; Length 1662;  
 Best Local Similarity 97.2%; Pred. No. 0;  
 Matches 168; Conservative 0; Mismatches 1; Indels 42; Gaps 2;

QY 175 AGGTTTGGTTTCAGAGCTCAAGATGGCTGACAGACGACCCCTGGGCAATCAC 234  
 Db 106 aggtttgtttcagagctcaagatggctgacagacccctgggcaatcac 165  
 QY 235 CCCTTCACAAATGGTGGCTCAGGACACGGGGCTGTGGTTACCTCTCTTCATGACA 294  
 Db 166 ccccttcacaaatggctgagctcagagcaccgggctgtgttaacctcttcatagaca 225  
 QY 295 CCCCTGGAGCTGTGAAGTTGGCTGACGATCAGCGGCCCTCCATGGCCAGCAGCTG 354  
 Db 226 cccctggagctgtgaagttggctgacgattcagcgccctccatgagcagagctg 285

QY 355 ATGCTTCCTCCAGACTGTGGAG-----CCTTCCTATAC 390  
 Db 286 atgcttctccacagctgtggagcctctccataacaaatgcccctccctccataac 345  
 QY 391 AAATGGAAGTGCCTCTCTATTGGCAATGTGTCTCGAGACCTCTTACTGTGCCCAAT 450  
 Db 346 aaatggaagtgcctctctattggcaatgtgtctcgagacctcttactgtgccaaat 405  
 QY 451 GGTGCCCGGTGGCCACTGGTTTCAGACCCCTTACCCGTTTCACTGGACCACTGATGCC 510  
 Db 406 ggtgcccggtggccactgggttcagaccccttaccgcttcaactggaactgagtc 465  
 QY 511 TTCTGTAAGATCGTAGGACACGAGGACACGACCCCTCTGGAGCGGCTCCCGCCAC 570  
 Db 466 ttctgtaagatcgttaggacacgagagccaccagacctcttgagagcgctcccgccac 525  
 QY 571 CTGGTGATGACTGTGCCAGACTACCGCATCTTACTTCACTAGCTTATGACCACTGAAGGC 630  
 Db 526 ctgtgtgactgtgtgccagactaccgcatcttacttcaactgtcctatgacccaactgagagcc 585  
 QY 631 TTCTGTGTGGTGGAGCCCTGACTCTGTACCTTACGACCCCATGTGCTGGCGGCTG 690  
 Db 586 ttctgtgtgtggaggccctgactctgtacttcaactgagctgtgagagctgtgc 645  
 QY 691 GCCCGCTTGGGACCGCTGACTGTGATCAACCCCTTGAGACTTATGCGCAAAAGCTGCAG 750  
 Db 646 gcccgcttgggacccgtgactgtgatcaaccccttgagacttatgcgcaaaagctgcag 705  
 QY 751 GCTCAGCATGTGTCTGTACCGGAGCTGGGTGCTGTTCGAACGTGCACTGAGGT 810  
 Db 706 gctcagcatgtgtctgtaccggagctgggtgctgttcgaactgagctgagctcagaggt 765  
 QY 811 GGCTGGCGCTCACTGTGGCTGGGGCTGGGGCCCGACCTGCGAGATGCGCTTCTC- 869  
 Db 766 ggctggcgctcactgtgtggctggggctggggcccgactgctcgagatgtgcctctc 825  
 QY 870 -----ACCGTGTACTGTTCAACTATGAGCTGGTGAAGAGCTGGCTC 912  
 Db 826 gtcatcccccaccccaagccctgtactgttcaactatgagctgtgagagctgtgc 885  
 QY 913 AATGGGCTCAGGCCCAAGAGACGACACTTCTGTGGGATGATGACTTGTGGCTGTGGCANC 972  
 Db 886 aatgggctcagggcccaagagacgacttctgtgggagctgtgtgtgtgtgtgtgtgtgt 945  
 QY 973 TCAGGGAGGTGAGCTGCACTGTGACTTACCTTTGAGCTGTAAAGCCCAAGCCAG 1032  
 Db 946 tcaaggaggtgagctgcaactgtgacttacccttgaacgtgttaagagcccaagccag 1005  
 QY 1033 GTGCTGTGGAGCGATGAGGCTGTGAGAGTGAACCCCTGCATGTGAGCTCCACTGG 1092  
 Db 1006 gtgcttggagcgatgagagctgtgagagtgagaaacccctgcatgtgagctccacactg 1065  
 QY 1093 CTGTGCTGCGAGGATCCGGGCCGAGTCCGACCAAGGAGACTCTTTGACGCTTCTC 1152  
 Db 1066 ctgtgctgcgagagatccggggccgagtcggacccaagagactcttgcagcttctc 1125  
 QY 1153 CCTCGATATCAAGGCTCCCTCTCTGCTGCTATCATATGACACCTATGATGTTGGC 1212  
 Db 1126 cctcgatatacaaggctccctctctgtgcatcatagtatgaacccctatgaagtctgcg 1185  
 QY 1213 AAAAGCTTCTCAGAGCTGAACAGGACCGGCTTCTGGGGGCTGAAGAGGGCAAGGA 1272  
 Db 1186 aaaagcttctcagagctgaacaggacccgcttctgagagctgaaaggggcaagga 1245  
 QY 1273 GGCAAGAGCCCGTCTCTCCACGAGTGGGAGAGGACGAGAGACCCAGCAAGTGC 1332  
 Db 1246 ggcaagagcccgctctctccacagatggggagagggcagggagagcccaagcgaatgc 1305  
 QY 1333 CTTTTCCTCAGACATGAGGAGAGGGGCTGTTTCTCTTCCCTCCCGGCAACAAGCTCAG 1392  
 Db 1306 ctttctcctaagacatgagagaggggctgttctctctccctcccgagcaagagctccag 1365  
 QY 1393 GGCAGGAGTGTCCCTGTGGGCGGCCAGACATTCCTCAAGACAACCTTCTCTGTGCT 1452

```
|||||
Db 1366 ggcagagcgtccctcctggtggcgccacgacctccctcagacaaactcttcctgtgct 1425
QY 1453 CCAGTCGTGGGATATCATCTACCCACCCCAAGTTCAACACCAATCTCCAGCTGC 1512
Db 1426 ccagtcgtggtgatcatctaccaccccaagttcaagaacaaatcttcagctgc 1485
QY 1513 CCCCTTCGTGTTCCCTGTTGCTGTAGTGGGCATGTCTCAGAACCAAGAGCCC 1572
Db 1466 ccccttcgtgttccctcctgtgttgcgtgtagctgggcaatgcccaagaagccc 1545
QY 1573 TCAGCCTGGTGTAGTCTCCCTGACCCCTGTGTAATTCCTTAAGTCAAGATGATGAATT 1632
Db 1546 tcagcctggtgtagctccctgcagccttgtaattccttaagtaagatgatgaact 1605
QY 1633 CAAAAAATAA 1643
Db 1606 caaaaaaaa 1616

RESULT 3
AAA06606
ID AAA06606 standard; cDNA: 537 BP.
XX
AC AAA06606;
XX
DT 13-JUN-2000 (first entry)
XX
DE Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:387.
XX
KW Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW immunogenic; cytosolic; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200004149-A2.
XX
PD 27-JAN-2000.
XX
PF 14-JUL-1999; 99WO-US15838.
XX
PR 14-JUL-1998; 98US-0115453.
PR 14-JUL-1998; 98US-0116134.
PR 23-SEP-1998; 98US-0159812.
PR 23-SEP-1998; 98US-0159822.
PR 15-JAN-1999; 99US-0232149.
PR 15-JAN-1999; 99US-0232880.
PR 09-APR-1999; 99US-0288946.
XX
FA (CORI-) CORIXA CORP.
XX
PI Dillon DC, Harlocker SL, Yugu J, Xu J, Mitcham JL;
XX
DR WPI: 2000-171268/15.
XX
PT New polypeptide useful for treating and diagnosing prostate cancer
PT comprises an immunogenic portion of prostate tumor protein -
XX
PS Claim 50; Page 235-236; 263pp; English.
XX
CC The present invention describes isolated polypeptides, comprising an
CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC and polynucleotides encoding them have cytostatic activity and can be
CC used in vaccines and in gene therapy. The polypeptides and
CC polynucleotides encoding them, antigen presenting cells which express
CC the polypeptides, antibodies against the polypeptides and vaccines
CC comprising them can be used for inhibiting the development of prostate
CC cancer in a patient. The polypeptides can be used to generate antibodies
CC or anti-idiotypic antibodies for passive immuno therapy. A portion of
CC the polynucleotides encoding the polypeptides can be used as a probe or
CC to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
CC AA82000 to AA82020 represent sequences used in the exemplification of
CC the present invention.
```

```
XX
SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
Query Match 32.4%; Score 532; DB 21; Length 537;
Best Local Similarity 100.0%; Pred. No. 6,76-125;
Matches 532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1112 GGGCGAGTCGGGACCCAGGAGACCTTTGACAGGCTTCCTCTCGGATCAACAGGCTG 1171
Db 1 999ccagtcggtgacacaaaggactcttgcaggtcttccttcgacatcaaaagctcg 60
QY 1172 CCCCTCTCTGTGCATCATGATCAGACACTATGATTCGGCAAAAGCTTCTTCAGAGGC 1231
Db 61 cccctcctcgtgcatcatcaatcagacactatgagttcgcaaaagcttcttcagagcc 120
QY 1232 TGAACAGGACCGGCTTCTGGCGGCTGAAGAGGGCAGAGGCAAGACCCGCTCTTC 1291
Db 121 tgaccagagccggtcttcgtggtcggtgaaagggtcaaggagcaagaccctctctc 180
QY 1292 CCACGATGGGAGAGGAGGAGAGAGACCAAGCCAGTCCCTTTCACAGCAGTGAAG 1351
Db 181 ccacggatggtggaagaggaaggaagaccagcaagtgcttcttcagcacatgagg 240
QY 1352 GAGGGGGCTTGTTCCTTCCTCCCGGACAGAGCTCCAGGCGAGGGCTGCTCTGG 1411
Db 241 gagggggtctgttcccttcctccctcccgagacagctccagggcagggctgtcccttcg 300
QY 1412 GCGGCCAGACATTTCTCTCAGACACAATTCTTCTGCTCTCAGTCTGGGGATCATCA 1471
Db 301 ggcgcacagactctcctcagacacactcttcctcgtcagtcgtygggatacata 360
QY 1472 CTTACCCACCCCCCAAGTTCAAGACCAAACTTCCAGCTGCCCTCGTGTTCCTGT 1531
Db 361 cttaaccacccccaagttaagaagacaaactctcagctgccttcgtgttccctcgt 420-
QY 1532 GTTTCGTAGCTGGGCACTGTCTCCAGAACCAAGAGCCCTCAGCTGGTGTAGTCTCC 1591
Db 421 gtttgtgtagctggtgcatgtctccaggaaccaagaagccctcagcctgtgtagctcc 480
QY 1592 CTGACCCCTGTGTAATTCCTTAAGTCAAGATGATGAACTTCAAAAAATAA 1643
Db 481 ctagacctgttaattccttaagtaagatgacttaaaaaa 532

RESULT 4
AAH93722
ID AAH93722 standard; cDNA: 537 BP.
XX
AC AAH93722;
XX
DT 04-OCT-2001 (first entry)
XX
DE Human prostate-specific cDNA sequence CGI-69.
XX
KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW cytosolic; gene therapy; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200151633-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US01574.
XX
PR 14-JAN-2000; 2000US-0483672.
XX
FA (CORI-) CORIXA CORP.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Rajos MD, Fanger GR, Day CH, Retter MM, Stolk JA, Skelky YAM;
PI Wang A, Meagher MT;
```



```

Oy 1412 GCGGCCAGACACTTCTCTCAGACACAACTTCTCTGCTGCTCCAGTGTGGGATCATCA 1471
Db 301 gggggccagcacttcctcctcagacacacttctcctgctgcagtcgctgggagatca 360
Oy 1472 CTTCACCAACCCCAAGTCAAGACCAAAATCTTCAGCTGCCCTTCTGTTTCCCTGT 1531
Db 361 cttaaccaccccccaagctcctcctcagacacaaatcttcagctgcgcccttcgtttccctgt 420
Oy 1532 GTTGTCTAGCTGGGCGATGTCCTCCAGGACCAAGAACCTTCAGCTGGGTAGTCTCC 1591
Db 421 gttgtctgtagctggtgcactgtctccaggaaccagaagccctcagcctggtgtagctcc 480
Oy 1592 CTGACCCCTGTAAATCTCCTTAAGTCTAAAGATGATGAACTTCAAAAAA 1643
Db 481 ctgaccctgttaattccttaagctaaagtgtgtaacttcaaaaaa 532

RESULT 6
ID AAH02787
XX AAH02787 standard; cDNA: 537 BP.
AC AAH02787;
XX
XX 14-JUN-2001 (first entry)
XX
XX Prostate tumour antigen cDNA sequence for CGI-69.
DE
XX
XX Human: prostate tumour antigen; prostate tumour; therapy; diagnosis;
KW prostate cancer; immunogenic; cytosolic; vaccine; ss.
XX
OS Homo sapiens.
XX
XX WO200125272-A2.
XX
XX 12-APR-2001.
XX
XX 04-OCT-2000; 2000WO-US27464.
XX
XX 04-OCT-1999; 99US-0157455.
XX
XX 04-OCT-1999; 99US-0157455.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Skeiky YAM, Reed SG, Cheever MA;
XX
XX WPI: 2001-245062/25.
XX
XX Prostate specific protein and its encoding polynucleotide, useful for
XX the treatment and diagnosis of prostate cancer -
XX
XX Claim 50; Page 244; 276pp; English.
XX
XX The present invention describes an isolated polypeptide (I) comprising
XX at least an immunogenic portion of a prostate tumour antigen protein or
XX its variant. (I) have cytostatic activity and can be used in vaccine
XX production. (I), prostate tumour antigen polynucleotides, an antigen
XX presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
XX pharmaceutical composition containing (I) are useful for inhibiting the
XX development of cancer in a patient. Antibodies specific for prostate
XX specific proteins and oligonucleotides that hybridise to a
XX polynucleotide that encodes a prostate specific protein are useful
XX for detecting the presence or absence of a cancer or monitoring the
XX progression the progression of a cancer, especially prostate cancer.
XX AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences
XX used in the exemplification of the present invention.
XX
XX Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
XX
XX
XX Query Match 32.4%; Score 532; DB 22; Length 537;
XX Best Local Similarity 100.0%; Pred. No. 6,7e-125;
XX Matches 532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1112 GGGCGCAGTGGGGCCACCAAGGACTTTCAGAGCTTCTTCCTCGGATCATCAAGCTG 1171

```

Dd		1	gagcgagtcggagacaaaggactcttgcaggtcttcctccctcgcgaalcaacaagtctg	60			
Oy	1172	CCCCCTCCTGTGGCCATCATGATCATGCAGCACCACTATGATTGGCAAAAGCTTTCACAGAGC		1231			
Dd	61	cacctctctgycatcatgatgtatgaaccctatgatttcgcgaagaagctctctccaagagc		120			
Oy	1232	TGACACGAGACC GGCTTGCGGGCGCTGGAAGGGGCAAGAGAGGCAAGGACCAAGGCCGTCCTC		1291			
Dd	121	tgaaccagagaccggctcttggtgcgtctgaaggagcaaggagcaaggaccgcgtctctc		180			
Oy	1292	CCACGATGGGGAGAGGGGCAAGAGAGAACCCAGCAAGTGCCTTTTCCTCAGCAGTAGAG	1351				
Dd	181	ccaagtaggggagagagagcagagagagagacccaagtagctcttcctcagacatgag	240				
Oy	1352	GAGGGGGCTTTGTTCCCTTCCTCCGCCGCGACAAGCTCCAGGGCAGGGCTGTCCCTCTGG	1411				
Dd	241	gaggggagtttgttccctctccctccgcgagcaagctccagggcagggctgctccctctg	300				
Oy	1412	GGGGCCGACACTTCCTCTAGACACAACTTCTCTGCTGCATCCAGTCGAGTGGATCATCA	1471				
Dd	301	ggcgcccagcaccttccctaagacaaactcttcctgctgccagtcgagtcgagatcata	360				
Oy	1472	CCTTACCACCCCCCAGTTCAGAAGCAAATACTTCTCACCTCCGCCCTTCTGATTCCCTGT	1531				
Dd	361	ctttaccaccccccaagttaaacagcaaatcttcacgtgcgcccttcggtctccctgt	420				
Oy	1532	GTTTGCCTGATGCTGGGCAATGTCCTCCAGAGACCAAGACCCTCAGCCTGGTGTAGTCTGC	1591				
Dd	421	gtttgctgtagctggcagctgctccaggaaccaaagaccctcagcctggtgtagttcc	480				
Oy	1592	CTGACCCCTGTGTAATTCCTTAAGTCCTAAGATGATGATGATGAACTTCAAAAAAA	1643				
Dd	481	ctgacctgtttaattccttaagttcaagatgatgaaacttcaaaaaaaaaa	532				
<hr/>							
RESULT 7							
ID	AAC76691/c						
XX	AAC76691 standard; cDNA; 6712 BP.						
XX							
DT	AAC76691;						
XX							
DE	08-FEB-2001 (first entry)						
XX							
XX	Human ORF246 polynucleotide sequence SEQ ID NO:4491.						
XX							
KW	Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;						
KW	vulnerary; antipoxviralic; antiparkinsonian; nootropic; neuroprotective;						
KW	anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;						
KW	immunostimulant; thrombolytic; coagulant; vasotrophic; antidiabetic;						
KW	hypotensive; dermatological; immunosuppressive; antinflamatory;						
KW	antiviral; antibacterial; antifungal; antirheumatic; antithyroid;						
KW	antianaemic; gene therapy; cancer; proliferative disorder; hypertension;						
KW	neurodegenerative disorder; osteoarthritis; graft vs host disease;						
KW	cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;						
KW	cholesterol ester storage; systemic lupus erythematosus; infection;						
KW	severe combined immunodeficiency; malaria; autoimmune disorder; asthma;						
KW	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;						
KW	bone damage; cartilage dsage; antinflamatory disease; coagulation;						
XX	thrombosis; contraceptive; ss.						
OS	Homo sapiens.						
XX							
PN	WO200058473-A2.						
XX							
PD	05-OCT-2000.						
XX							
PF	31-MAR-2000; 2000MO-US08621.						
XX							
PR	31-MAR-1999; 99US-0127607.						
PR	02-APR-1999; 99US-0127636.						
PR	05-APR-1999; 99US-0127728.						
PR	30-MAR-2000; 2000US-0540763.						



CC sequences were obtained from 3'-directed cDNA libraries prepared  
CC from various human tissues; synthesis of cDNA was initiated from the  
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
CC untranslated sequence is unique to a particular mRNA species, almost  
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library  
CC is constructed so as to reflect accurately the relative abundance of  
CC different mRNAs in the particular tissue from which it was derived.  
CC The appearance frequency of a given GS in a cDNA library can be  
CC determined (esp. using primers and probes derived from the GS  
CC sequences) as a means of diagnosing abnormal cell function or for  
CC recognising different cell types.

SQ Sequence 174 BP; 40 A; 52 C; 30 G; 48 T; 4 other;

Query Match 10.2%; Score 168.4; DB 16; Length 174;  
Best Local Similarity 97.1%; Pred. No. 3e-33;  
Matches 169; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1464 GATCATCACTTACCCACCCCAAGTTCAGACAAATCTTCAGCTGCCCTTCGTGT 1523  
|||||  
DB 1 gatcatcacttaccaccccccaagttcaagaacaaatcttcagctgcccttcgtgt 60  
OY 1524 TTCCTGTGTGTGTGTAGCTGGGCATGCTTCAGAACCAAGAACCCCTAGCCTGTG 1583  
|||||  
DB 61 ttccctgt 120  
OY 1584 TAGTCTCCCTGACCTTGTATTCCTTAAGTCTAAAGATGATGAACTTCAAAA 1637  
|||||  
DB 121 tagtctccctgacctttnaatcttcaagtcataagatgatacttctaa 174

RESULT 9  
AAH30252  
ID AAH30252 standard; cDNA: 397 BP.  
XX  
AC AAH30252;  
XX  
DT 27-JUL-2001 (first entry)  
XX  
DE Human colon cancer cell line Km12L4-A cDNA library derived sequence #186.  
XX  
KW Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;  
KW detection; colon cancer cell line Km12L4-A; ss.  
XX  
OS Homo sapiens.  
XX  
PM WO200018916-A2.  
XX  
PD 06-APR-2000.  
XX  
PF 23-SEP-1999; 99WO-US22226.  
XX  
PR 28-SEP-1998; 98US-0102161.  
PR 28-SEP-1998; 98US-0102180.  
PR 29-SEP-1998; 98US-0102380.  
PR 08-OCT-1998; 98US-0103815.  
PR 27-OCT-1998; 98US-0105877.  
XX  
PA (CHIR ) CHIRON CORP.  
PA (HYSE-) HYSEQ INC.  
XX  
PI Williams JT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;  
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;  
PI Lanson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;  
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;  
XX  
DR WPI; 2000-293155/25.  
XX  
XX Polynucleotide library comprising 1079 defined sequences, useful in  
PT the form of an array to detect cancer or susceptibility to cancer -  
XX  
PS Claim 1; Page 23; 502PP; English.

XX The present invention describes a library of polynucleotides comprising  
CC 1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described  
CC are: (1) an isolated polynucleotide (I) having at least 90% identity to  
CC one of the 1079 sequences; (2) a recombinant host cell containing (I);  
CC (3) an isolated polypeptide (II) encoded by (I); (4) an antibody that  
CC specifically binds to (II); (5) a vector comprising (I); and (6) a method  
CC of detecting differentially expressed genes correlated with a cancerous  
CC state of a mammalian cell comprising detecting a gene product encoded by  
CC one of the 1079 sequences given in the specification. The polynucleotides  
CC are used to monitor patients having (or susceptible) to cancer to detect  
CC potentially malignant events at a molecular level before they are  
CC detectable at a gross morphological level. The polynucleotides are also  
CC useful for monitoring the efficacy of various therapies and preventive  
CC interventions. Polynucleotide probes based on the disclosed sequences  
CC are useful for chromosome mapping and detection of transcription levels.  
CC The 1079 polynucleotide sequences were derived from a human colon cancer  
CC cell line Km12L4-A cDNA library.

SQ Sequence 397 BP; 110 A; 70 C; 90 G; 127 T; 0 other;

Query Match 8.8%; Score 144.2; DB 21; Length 397;  
Best Local Similarity 62.2%; Pred. No. 5.1e-27;  
Matches 227; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

OY 545 CCCCTGGAGCGGCGCTCCCGCCACCTGTGATGATGACTGTGCCACTACCGCATCTACT 604  
|||||  
DB 4 cctctagagtggtccttctctctaccctagtgatgagtcagttcttcacagttattt 63  
OY 605 TCACATGCCCTATGACCACTGAAGGCTTCCTGTGTGTGTGTGTGTGTGTGTGTGT 664  
|||||  
DB 64 ttactctgcatgataatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 123  
OY 665 ACGCAACCATGCTGGCTGGCGGCTGGCCCGCTTGGGCAACCGTGAAGTGAAGTGAAGT 724  
|||||  
DB 124 gataccaattgtgtgtgaattgtgaattgtgaattgtgaattgtgaattgtgaattgtga 183  
OY 725 TGGAGCTTATGCGGACCAAGCTGACAGGCTACACATGTGTGTGACCGGAGCTGGGTGCT 784  
|||||  
DB 184 tgaattgattgataccaagaatgacagtcacgaatgatttcttaagtgtaactgcatgat 243  
OY 785 GTTTCGAATGCTAGTGTGCTGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 844  
|||||  
DB 244 ttgtcagaagaagaatgataatgataatgataatgataatgataatgataatgataat 303  
OY 845 CTGCCCTCGAGATGTCGCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 904  
|||||  
DB 304 ctgtccttagagatgtaccttcttcagcaatgacatgacatgacatgacatgacatgacat 363  
OY 905 GCTGG 909  
DB 364 agtgg 368

RESULT 10  
AAI39053/C  
ID AAI39053 standard; DNA: 452 BP.  
XX  
AC AAI39053;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Probe #7739 used to measure gene expression in human placenta sample.  
DE Probe; microarray; human; placenta; antenatal diagnosis;  
XX  
KW genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PM WO200157272-A2.  
XX  
PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488897/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human placenta -  
 XX  
 PS Claim 25; SEQ ID No 7739; 654pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SENPs).  
 CC The present sequence is one such probe. The probes are useful for  
 CC producing a microarray for predicting, measuring and displaying gene  
 CC expression in samples derived from human placenta. The probes are useful  
 CC for antenatal diagnosis of human genetic disorders.  
 XX  
 SQ Sequence 452 BP; 181 A; 83 C; 72 G; 116 T; 0 other;

Query Match 4.6%; Score 74.8; DB 22; Length 452;  
 Best Local Similarity 64.4%; Pred. No. 1.8e-09;  
 Matches 112; Conservative 0; Mismatches 62; Indels 0; Gaps 0;  
 QY 698 TGGGCACCGTACTGTGTATGACAGCCCTGAGAGCTTATGCGGACAAAGCTGACAGGCTCAGC 757  
 DB 436 TTGGTGACAGTACTGTATGATTAAGTCAAGATTGATTAAGAACCAAGATGACAGTCCAGCA 377  
 QY 758 ATGTGTCTACCGGAGAGCTGGGCTGTGTTCCAGACTGCAGTGCAGGCTGAGGCTGGC 817  
 DB 376 AGTTTCTTACGTGAGACTGCATTGTTCAGCAGAAAGATATCTAAGATGGTTGGA 317  
 QY 818 GCTCAGCTGGCTGGCTGGGAGCCCACTGCGCTTCAGAGATGGCCCTTCAG 871  
 DB 316 TTTCCTTTGAGAGGGGCTGGGCTCTCTACTGTTCCTTAGAGATGTACCTTTCTCAG 263

RESULT 11  
 AA139123  
 ID AA139123 standard; DNA; 496 BP.  
 XX  
 AC AA139123;  
 XX  
 DT 17-OCT-2001 (first entry)  
 XX  
 DE Probe #7809 used to measure gene expression in human placenta sample.  
 XX  
 KW Probe: microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157272-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 XX

PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488897/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human placenta -  
 XX  
 PS Claim 25; SEQ ID No 7809; 654pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SENPs).  
 CC The present sequence is one such probe. The probes are useful for  
 CC producing a microarray for predicting, measuring and displaying gene  
 CC expression in samples derived from human placenta. The probes are useful  
 CC for antenatal diagnosis of human genetic disorders.  
 XX  
 SQ Sequence 496 BP; 150 A; 65 C; 95 G; 186 T; 0 other;

Query Match 4.6%; Score 74.8; DB 22; Length 496;  
 Best Local Similarity 64.4%; Pred. No. 1.8e-09;  
 Matches 112; Conservative 0; Mismatches 62; Indels 0; Gaps 0;  
 QY 698 TGGGCACCGTACTGTGTATGACAGCCCTGAGAGCTTATGCGGACAAAGCTGACAGGCTCAGC 757  
 DB 259 ttggtgcagtaactgtgtaagtcacactagaattgattagaaccaagaatgcaagtcacaaga 318  
 QY 758 ATGTGTCTACCGGAGAGCTGGGCTGTGTTCCAGACTGCAGTGCAGGCTGAGGCTGGC 817  
 DB 319 agtttcttaagtggaactgcacgattgtcagaagaagaatctatctgaagaatggtctgga 378  
 QY 818 GCTCAGCTGGCTGGCTGGGAGCCCACTGCGCTTCAGAGATGGCCCTTCAG 871  
 DB 379 ttcccttggaggggctgggctctctactgttctttagagaagtacttctcag 432

RESULT 12  
 AA210752/C  
 ID AA210752 standard; DNA; 72604 BP.  
 XX  
 AC AA210752;  
 XX  
 DT 23-NOV-1999 (first entry)  
 XX  
 DE Genomic sequence of the human HKNG1 gene.  
 XX  
 KW HKNG1; Hong Kong new gene 1; bipolar affective disorder; BAD;  
 KW neuropsychiatric disorder; early-onset autosomal dominant myopia;  
 KW schizophrenia; splice variant; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9947535-A1.  
 XX  
 PD 23-SEP-1999.  
 XX  
 PF 16-MAR-1999; 99WO-US05606.  
 XX  
 PR 16-MAR-1998; 98US-0078044.  
 PR 05-JUN-1998; 98US-0088312.  
 PR 28-OCT-1998; 98US-0106056.  
 PR 22-JAN-1999; 99US-0236134.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Chen H, Freimer NB;  
 XX

DR WPI: 1999-562047/47.  
XX New HKNG1 polynucleotides useful in diagnosis and treatment of  
PT neuropsychiatric disorders, e.g. bipolar affective disorders and  
PT schizophrenia  
XX  
PS Claim 2; Fig 3A-R; 205pp; English.  
XX  
CC The present sequence represents HKNG1 (Hong Kong new gene 1). HKNG1  
CC is a gene associated with bipolar affective disorder (BAD). HKNG1  
CC polynucleotides are useful to identify compounds modulating HKNG1  
CC expression or HKNG1 polypeptide expression/activity. Compounds inhibiting  
CC or enhancing HKNG1 gene expression or activity in individuals can then  
CC be administered therapeutically to treat HKNG1-mediated disorders,  
CC especially neuropsychiatric disorders e.g. BAD, schizophrenia, or  
CC HKNG1-mediated myopia disorders, such as early-onset autosomal  
CC dominant myopia. The polynucleotides can be used in gene therapy  
CC techniques to treat such disorders. They are also useful in diagnosis  
CC to identify individuals having, or at risk of developing, HKNG1-mediated  
CC disorders due to mutations in the HKNG1 gene. Such mutations especially  
CC result in the production of a protein with a different sequence to  
CC the human full-length HKNG1 polypeptide or splice variant sequences,  
CC especially the substitution of a lysine for a glutamic acid at residue  
CC 202 or 184. The polynucleotides are also useful in gene mapping, to  
CC produce probes or primers to identify similar sequences (e.g. mutants  
CC or sequences from different species) and to produce transgenic  
CC animals.  
XX  
SQ Sequence 72604 BP; 20579 A; 15146 C; 14859 G; 21900 T; 120 other;  
Query Match  
Best Local Similarity 4.0%; Score 65.4; DB 20; Length 72604;  
Matches 100; Conservative 1; Mismatches 12; Indels 8; Gaps 2;  
OY 2 GGAGGCTAAGCAGAGAACCGCTTGAAAC---GGGGGTGATGTTGACGTGACCAAGAT 58  
DB 9317 GGAGGCTAAGCAGAGAACCGCTTGAAACCTGGGAGGTGGAGTTCAGTGGCCAAAGAT 9258  
OY 59 GGTGCTACTGACCTGCACCTGGCAAGAGTGGCAGAGAGAGCCGCTCAAAACAA 118  
DB 9257 CGTGCACTGCACCTGCACCTGG---GTGAGCAGAGCAGACATCTCAAAACAA 9203  
OY 119 A 119  
DB 9202 A 9202  
RESULT 13  
AAC77105/c  
ID AAC77105 standard; cDNA: 4231 BP.  
XX  
AC AAC77105;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Human ORFX ORF2660 polynucleotide sequence SEQ ID NO:5319.  
XX  
XX Human ORFX ORF2660 polynucleotide sequence SEQ ID NO:5319.  
XX  
KM Human: open reading frame; ORFX: detection; cytosolic; hepatotropic;  
KM vulnary; antipsoriatic; antiparkinsonian; noctropic; neuroprotective;  
KM anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardial;  
KM hypotensive; dermatological; immunosuppressive; antidiabetic;  
KM antiviral; antibacterial; antifungal; antineoplastic; antihypertensive;  
KM antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
KM neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KM cholesterol ester storage; systemic lupus erythematosus; infection;  
KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KM bone damage; cartilage damage; antiinflammatory disease; coagulation;  
KM thrombosis; contraceptive; ss.  
XX

OS Homo sapiens.  
XX  
XX W0200058473-A2.  
XX  
XX 05-OCT-2000.  
XX  
XX 31-MAR-2000; 2000WO-US08621.  
XX  
XX 31-MAR-1999; 99US-0127607.  
XX 02-APR-1999; 99US-0127636.  
XX 05-APR-1999; 99US-0127728.  
XX 30-MAR-2000; 2000US-0540763.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shinkets RA, Leach M;  
XX  
XX WPI: 2000-602362/57.  
XX P-PSDB: AAB42896.  
XX  
XX Novel nucleic acids and peptides derived from open reading frame X,  
XX useful for treating e.g. cancers, proliferative disorders,  
XX neurodegenerative disorders and cardiovascular disease -  
XX  
XX Claim 5; Page 4490-4493; 5507pp; English.  
XX  
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX  
XX sequences have activities such as: cytosolic; hepatotropic; vulnary;  
XX antipsoriatic; antiparkinsonian; noctropic; neuroprotective;  
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;  
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
XX antidiabetic; hypotensive; dermatological; immunosuppressive;  
XX antiinflammatory; antibacterial; antiviral; antifungal; antineoplastic;  
XX antihypertensive; antianaemic. The sequences can be used for determining  
XX the presence of or predisposition to, or preventing or treating  
XX pathological conditions associated with an ORFX-associated disorder. The  
XX nucleic acids can be used to express ORFX proteins in gene therapy  
XX vectors. The proteins and nucleic acids may be used to treat cancers,  
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,  
XX graft vs host disease, cardiovascular disease, diabetes mellitus,  
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
XX coagulation; to inhibit thrombosis; and as a contraceptive.  
XX  
XX Sequence 4231 BP; 832 A; 1348 C; 1095 G; 953 T; 3 other;  
SQ  
Query Match  
Best Local Similarity 3.9%; Score 64.2; DB 21; Length 4231;  
Matches 100; Conservative 0; Mismatches 14; Indels 8; Gaps 2;  
OY 1 GGGAGCTGAAGCAGAGAACCGCTTGAAAC---GGGGGTGATGTTGACGTGACCAAGA 57  
DB 3131 GGGAGCTGAAGCAGAGAACCGCTTGAAACCTGGGAGGTGGAGTTCAGTGGCCAAAGA 3072  
OY 58 TGTGCTACTGACCTGCACCTGGCAAGAGTGGCAGAGAGAGCCGCTCAAAACAA 117  
DB 3071 TGTGCTACTGACCTGCACCTGG---GTGAGCAGAGCAGACATCTCAAAACAA 3017  
OY 118 AA 119  
DB 3016 AA 3015  
RESULT 14  
AAF65290  
ID AAF65290 standard; cDNA: 444 BP.  
XX  
AC AAF65290;  
XX



Sat Apr 20 10:03:16 2002

us-09-840-787-68.rng

Page 12

Db 53 TTGTGCACCTGCACCTGGGCGCCACGACGAGACTCTGTCTC 1

Search completed: April 19, 2002, 21:45:56  
Job time: 4270 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2002, 19:42:21 ; Search time 88.57 Seconds  
(without alignments)  
4201.235 Million cell updates/sec

Title: US-09-840-787-68  
Perfect score: 1643  
Sequence: 1 GGGAGCGCTGACGACGAGAA.....GATGACTTCAAAAA 1643

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/2/1na/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/1na/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/1na/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/1na/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/1na/PCRTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/1na/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1643	100.0	1643	2	US-08-933-750C-68
2	1643	100.0	1643	3	US-09-234-613-68
3	62.8	3.8	87350	3	US-08-781-891-79
4	62	3.8	5009	3	US-08-978-741-7
5	62	3.8	5009	4	US-09-333-729A-8
6	61	3.7	997	3	US-09-188-930-23
7	61	3.7	1816	3	US-09-188-930-262
8	60.2	3.7	591	4	US-09-385-982-406
9	60.2	3.7	14636	4	US-09-173-914-6
10	59.6	3.6	330	3	US-09-157-177-114
11	59	3.6	176373	3	US-09-128-155-17
12	58.6	3.6	5137	5	PCT-US96-01314-39
13	58.6	3.6	5138	2	US-08-476-062A-39
14	58.4	3.6	2115	1	US-08-395-800A-7
15	58.4	3.6	19011	1	US-08-310-356-36
16	58.4	3.6	19557	1	PCT-US92-06300-1
17	58.4	3.6	246240	2	US-08-724-394A-20
18	58.4	3.6	246240	2	US-08-724-394A-21
19	58.4	3.6	246240	2	US-08-724-394A-22
20	58	3.5	20303	1	US-08-370-875B-6
21	58	3.5	26764	1	US-08-370-875B-1
22	57.8	3.5	352	4	US-09-385-982-442
23	57.8	3.5	606	4	US-09-385-982-413
24	57.8	3.5	632	4	US-09-385-982-177
25	57.8	3.5	657	4	US-09-385-982-335
26	57.8	3.5	1278	4	US-08-909-965C-4
27	57.8	3.5	2839	3	US-08-468-856B-5

28	57.8	3.5	2839	3	US-08-468-859A-5	Sequence 5, Appl1
29	57.6	3.5	498	4	US-09-085-189B-36	Sequence 36, Appl1
30	57.6	3.5	1554	1	US-08-370-975B-10	Sequence 10, Appl1
31	57.4	3.5	1656	1	US-08-324-465-2	Sequence 2, Appl1
32	57.4	3.5	1656	2	US-08-465-981-2	Sequence 2, Appl1
33	57.4	3.5	1656	5	PCT-US93-11915-2	Sequence 2, Appl1
34	57.4	3.5	1725	1	US-08-324-465-5	Sequence 5, Appl1
35	57.4	3.5	1725	2	US-08-465-981-5	Sequence 5, Appl1
36	57.4	3.5	1725	5	PCT-US93-11915-5	Sequence 5, Appl1
37	57.4	3.5	3804	2	US-08-483-488-5	Sequence 5, Appl1
38	57.4	3.5	9837	1	US-08-832-883-68	Sequence 68, Appl1
39	57.4	3.5	9837	2	US-08-832-877-68	Sequence 68, Appl1
40	57.4	3.5	13158	2	US-08-687-080-105	Sequence 105, Appl1
41	57.2	3.5	7720	4	US-09-318-448-5	Sequence 5, Appl1
42	57.2	3.5	18443	4	US-08-078-284-6	Sequence 6, Appl1
43	56.8	3.5	1200	4	US-09-018-584A-37	Sequence 37, Appl1
44	56.8	3.5	5375	3	US-08-757-223-7	Sequence 7, Appl1
45	56.8	3.5	246240	2	US-08-724-394A-20	Sequence 20, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-933-750C-68  
Sequence 68, Application US/08933750C  
Patent No. 5932442  
GENERAL INFORMATION:  
APPLICANT: Lal, Preethi  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Shah, Purni  
APPLICANT: Au-Young, Janice  
APPLICANT: Yue, Henry  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/933,750C  
FILING DATE: September 23, 1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TEXT:  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1643 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: SYNOCAT01

CLONE: 724157  
US-08-933-750C-68

Query Match 100.0%; Score 1643; DB 2; Length 1643;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGGAGGCTGAAGCAAGAACCGCTTGAACGGGGGTGATGTTGATGAGCCAAAGATG 60
DB 1 GGGAGGCTGAAGCAAGAACCGCTTGAACGGGGGTGATGTTGATGAGCCAAAGATG 60
QY 61 TGCCATGCTACCTCCAGCTGGCAAGTGGCGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 61 TGCCATGCTACCTCCAGCTGGCAAGTGGCGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 121 TCTCAGATGAGCCAGCCCTCTTCTCCATCCCTGAGAGTTGGTCTGTTGAGTTT 180
DB 121 TCTCAGATGAGCCAGCCCTCTTCTCCATCCCTGAGAGTTGGTCTGTTGAGTTT 180
QY 181 GGTTCAGAGACTGAAGCTTCAAGATGCTGACAGAGACCTGCGGGCATGACCCCTC 240
DB 181 GGTTCAGAGACTGAAGCTTCAAGATGCTGACAGAGACCTGCGGGCATGACCCCTC 240
QY 241 CAGCAAAATGATGGCTCAGGACCGGGGCTGTGTTACTCTCTTCATGACACCCCTG 300
DB 241 CAGCAAAATGATGGCTCAGGACCGGGGCTGTGTTACTCTCTTCATGACACCCCTG 300
QY 301 GACGATGTAAGGTTGCGCTGCAAGTCTGACGGCCCTCCATGGCCAGAGCTGATGCT 360
DB 301 GACGATGTAAGGTTGCGCTGCAAGTCTGACGGCCCTCCATGGCCAGAGCTGATGCT 360
QY 361 TCCCTCAGAGCTGAGAGCTCTCTCTATACCAATGGAAGTGGCTCTGATTTGCAATG 420
DB 361 TCCCTCAGAGCTGAGAGCTCTCTCTATACCAATGGAAGTGGCTCTGATTTGCAATG 420
QY 421 GTCTGAGAGCTCTGTACTGTGCCCCAATAGTGTGCCCGCTGTGCCACCTGTTTCAAG 480
DB 421 GTCTGAGAGCTCTGTACTGTGCCCCAATAGTGTGCCCGCTGTGCCACCTGTTTCAAG 480
QY 481 CCTACCCGCTTACTGTGACACCATGATGCTTCTGTGAAGATGTGAGGACAGAGGAC 540
DB 481 CCTACCCGCTTACTGTGACACCATGATGCTTCTGTGAAGATGTGAGGACAGAGGAC 540
QY 541 AGGACCTTGTGAGAGGCGCTCCCGCAACCTGTGATGACTGTGCACACTACCCCATC 600
DB 541 AGGACCTTGTGAGAGGCGCTCCCGCAACCTGTGATGACTGTGCACACTACCCCATC 600
QY 601 TACTTACTGCTATGACCACTGAAGGCTTCTGTGTGTGTCAGACCTGACTCTGAC 660
DB 601 TACTTACTGCTATGACCACTGAAGGCTTCTGTGTGTGTCAGACCTGACTCTGAC 660
QY 661 CTCTACGACCATGATGAGCAAGCTGAGAGGCTTCTGTGTGTGTCAGACCTGACTCTG 720
DB 661 CTCTACGACCATGATGAGCAAGCTGAGAGGCTTCTGTGTGTGTCAGACCTGACTCTG 720
QY 721 CCCCTGAGACTTATGCGCAAAAGCTGACAGCTGACATGTGCTGACCGGAGCTGGT 780
DB 721 CCCCTGAGACTTATGCGCAAAAGCTGACAGCTGACATGTGCTGACCGGAGCTGGT 780
QY 781 GCTGTGTTGCAAGCTGACATGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTG 840
DB 781 GCTGTGTTGCAAGCTGACATGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTG 840
QY 841 CCCAGTGCCTTCAGATGATGCTTCTGAGCCCTGACTGATGTTCAACTATGAGTGTG 900
DB 841 CCCAGTGCCTTCAGATGATGCTTCTGAGCCCTGACTGATGTTCAACTATGAGTGTG 900
QY 901 AAGAGCTGCTCAATGAGGCTCAAGGCGGAGAGACCAAGACTTCTGTGGGCAATGAGCTT 960
DB 901 AAGAGCTGCTCAATGAGGCTCAAGGCGGAGAGACCAAGACTTCTGTGGGCAATGAGCTT 960
QY 961 GCTGTGCTATCTCAGGAGAGGCTGCTGAGTGTGACTTACCTTTAGCGTGAAG 1020

```

```

DB 961 GCTGTGCTATCTCAGGAGAGGCTGCTGACTGCTGACTTACCTTTGAGCTGTAAG 1020
QY 1021 ACCCAACCCGAGCTGCTCTGTGGAGCGATGAGAGCTGTGAAGATGAAACCCCTGATG 1080
DB 1021 ACCCAACCCGAGCTGCTCTGTGGAGCGATGAGAGCTGTGAAGATGAAACCCCTGATG 1080
QY 1081 GACTTCACCTGCTGCTGCTGCTGAGAGATCCGGGCGGAGTGGGACCAAGGACTCTTT 1140
DB 1081 GACTTCACCTGCTGCTGCTGCTGAGAGATCCGGGCGGAGTGGGACCAAGGACTCTTT 1140
QY 1141 GCAGGCTTCTCTCTGAGATCATCAAGGCTGCCCCCTCTGTGCATCATATGACAC 1200
DB 1141 GCAGGCTTCTCTCTGAGATCATCAAGGCTGCCCCCTCTGTGCATCATATGACAC 1200
QY 1201 TATGATTCGCAAAAGCTTCTTCCAGAGGCTGAAACCAAGAGACCGGCTTGGGGGCTGA 1260
DB 1201 TATGATTCGCAAAAGCTTCTTCCAGAGGCTGAAACCAAGAGACCGGCTTGGGGGCTGA 1260
QY 1261 AAGGGCAAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 AAGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 CCAGGCAAGTGCCTTTCTCAGACACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 1321 CCAGGCAAGTGCCTTTCTCAGACACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 GACAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 1381 GACAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 CTTCCTGCTGCTCAGTGTGAGGAGATCATCACTTACCAACCCCAAGTTCAAGACAA 1500
DB 1441 CTTCCTGCTGCTCAGTGTGAGGAGATCATCACTTACCAACCCCAAGTTCAAGACAA 1500
QY 1501 TCTTCACGCTGCCCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
DB 1501 TCTTCACGCTGCCCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
QY 1561 ACCAAGAGCCCTCAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
DB 1561 ACCAAGAGCCCTCAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
QY 1621 GATGATGAACCTTCAAAAAAAAA 1643
DB 1621 GATGATGAACCTTCAAAAAAAAA 1643

```

# RESULT 2

US-09-234-613-68  
Sequence 68, Application US/09234613  
Patent No. 6132973

## GENERAL INFORMATION:

APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Shah, Purvi  
APPLICANT: Ah-Young, Janice  
APPLICANT: Yue, Henry  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible

```

: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/234.613
: FILING DATE:
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US/08/933.750
: FILING DATE: September 23, 1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0356 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
:
: INFORMATION FOR SEQ ID NO: 68:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1643 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: SYNDOT01
: CLONE: 724157
:
: US-09-234-613-68

```

```

Query Match          100.0%; Score 1643; DB 3; Length 1643;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 GGGAGGCTGAACAGAGAACCGCTTGAACGGGGGTGGATGTTGACAGTACGCAAGATGG 60
Db 1 GGGAGGCTGAACAGAGAACCGCTTGAACGGGGGTGGATGTTGACAGTACGCAAGATGG 60
Qy 61 TGCCACTGCATCCAGCCCTGCGACAGTGGCGAGACGAGCGCGCTCAAAACAAAAG 120
Db 61 TGCCACTGCATCCAGCCCTGCGACAGTGGCGAGACGAGCGCGCTCAAAACAAAAG 120
Qy 121 TCTCAGAGTGGCCAGCCCTCTTCTCTGCTGCAATCCCTGGAGTTGTGGTCTGTTGAGGTTT 180
Db 121 TCTCAGAGTGGCCAGCCCTCTTCTCTGCTGCAATCCCTGGAGTTGTGGTCTGTTGAGGTTT 180
Qy 181 GGTTCAGAGTGAAGCTTCAAGATGGCTGACAGAGACCCCTGGGGGCTACGCCCTGC 240
Db 181 GGTTCAGAGTGAAGCTTCAAGATGGCTGACAGAGACCCCTGGGGGCTACGCCCTGC 240
Qy 241 CAGCAAAATGATGGCTCTCAGAGCACGGGGGCTGTTTACCTCTCTCTTCATGACACCCCTG 300
Db 241 CAGCAAAATGATGGCTCTCAGAGCACGGGGGCTGTTTACCTCTCTCTTCATGACACCCCTG 300
Qy 301 GACGTGTGAAGGTTGCGCTGCAATGCTCAGCGGCCCTCATGGCAGAGCTGATGCTT 360
Db 301 GACGTGTGAAGGTTGCGCTGCAATGCTCAGCGGCCCTCATGGCAGAGCTGATGCTT 360
Qy 361 TCCCTCAAGCTGTGAAGCTCTCCCTATACCAATGGAATGGCTCTCTGATATGCAATGGT 420
Db 361 TCCCTCAAGCTGTGAAGCTCTCCCTATACCAATGGAATGGCTCTCTGATATGCAATGGT 420
Qy 421 GTCTGAGAGCTCTGTACCTGTGCCCCAAATGTTGCGCGCTGTCACACCGTGGTTTCAAGAC 480
Db 421 GTCTGAGAGCTCTGTACCTGTGCCCCAAATGTTGCGCGCTGTCACACCGTGGTTTCAAGAC 480
Qy 481 CCTACCCCGCTTCACTGAGCAGCATGATGCTTGTGAAGATCGTGAAGCAGAGGAGGAGG 540
Db 481 CCTACCCCGCTTCACTGAGCAGCATGATGCTTGTGAAGATCGTGAAGCAGAGGAGGAGG 540
Qy 541 AGGACCCCTCTGAGAGGCGCTCCCGGCAACCCCTGGTGAATGACCTGTGACCACTACGCCCATC 600
Db 541 AGGACCCCTCTGAGAGGCGCTCCCGGCAACCCCTGGTGAATGACCTGTGACCACTACGCCCATC 600

```

```

Qy 601 TACTTCATGCTCTATGACCACTGAAGGCTTCTGTGTGTGTCGAGCCCTGCATCTGCAC 660
Db 601 TACTTCATGCTCTATGACCACTGAAGGCTTCTGTGTGTGTCGAGCCCTGCATCTGCAC 660
Qy 661 CTCTACGACCCATGATGGCTGGCGGCTGGCCGCTTGGGACGCTGATGATCAGC 720
Db 661 CTCTACGACCCATGATGGCTGGCGGCTGGCCGCTTGGGACGCTGATGATCAGC 720
Qy 721 CCCCTGAGACTTATGCGGCAAAAGCTGACGGCTGACAGATGTCGTAACGGGAGCTGGGT 780
Db 721 CCCCTGAGACTTATGCGGCAAAAGCTGACGGCTGACAGATGTCGTAACGGGAGCTGGGT 780
Qy 781 GCGTGTGTCGACATGCTACATGCTCAGGGTGGCTGGCGCTCACTGTGCTGGGCTGGGCG 840
Db 781 GCGTGTGTCGACATGCTACATGCTCAGGGTGGCTGGCGCTCACTGTGCTGGGCTGGGCG 840
Qy 841 CCCACTGCCCTTTCAGAGATGTGCCCTTCTCAGCCCTGTACTGTGTTCAATATGACTGTG 900
Db 841 CCCACTGCCCTTTCAGAGATGTGCCCTTCTCAGCCCTGTACTGTGTTCAATATGACTGTG 900
Qy 901 AAGAGCTGGCTGAATGGGCTCAGGGCGGAGGACGAGCTTGTGGGAGTGAAGCTTGTG 960
Db 901 AAGAGCTGGCTGAATGGGCTCAGGGCGGAGGACGAGCTTGTGGGAGTGAAGCTTGTG 960
Qy 961 GCTGTGGCTACTCAGAGGAGGCTGGCTGCACTGCTACTCTACCTTGTGACGTGTAAG 1020
Db 961 GCTGTGGCTACTCAGAGGAGGCTGGCTGCACTGCTACTCTACCTTGTGACGTGTAAG 1020
Qy 1021 ACCCAAGCCGAGGTGCTCTGGGAGCGATGAGGCTGTGAGATGAAACCCCTGCATGTG 1080
Db 1021 ACCCAAGCCGAGGTGCTCTGGGAGCGATGAGGCTGTGAGATGAAACCCCTGCATGTG 1080
Qy 1081 GACTCCACCTGCTGCTGCTGGGAGATCCGGGCGAGATCGGGGACCAAGGAGACTCTT 1140
Db 1081 GACTCCACCTGCTGCTGCTGGGAGATCCGGGCGAGATCGGGGACCAAGGAGACTCTT 1140
Qy 1141 GCAGGCTTCTCTCTCTGATCATCAAGGCTGCCCCCTCTGTGCATCATGATGATGACAC 1200
Db 1141 GCAGGCTTCTCTCTCTGATCATCAAGGCTGCCCCCTCTGTGCATCATGATGATGACAC 1200
Qy 1201 TATGATTCGGCAAAAGCTTCTTCCAGAGGCTGAAACCAAGACCGGCTTGTGGCGGCTGA 1260
Db 1201 TATGATTCGGCAAAAGCTTCTTCCAGAGGCTGAAACCAAGACCGGCTTGTGGCGGCTGA 1260
Qy 1261 AAGGGGCAAGGAGGAGGACCCGCTCTCCACAGGATGGGAGAGGAGGAGGAGGAGC 1320
Db 1261 AAGGGGCAAGGAGGAGGACCCGCTCTCTCCACAGGATGGGAGAGGAGGAGGAGGAGC 1320
Qy 1321 CCAGCCAAAGTGCCTTCTCTCAGCACTGAGGAGGAGGCTTTCCTTCCCTCCCGGAC 1380
Db 1321 CCAGCCAAAGTGCCTTCTCTCAGCACTGAGGAGGAGGCTTTCCTTCCCTCCCGGAC 1380
Qy 1381 GACAAAGCTTCAGAGGAGGCTGCTCCCTGCGGCGCCAGCACTTCTCAGACACAACTT 1440
Db 1381 GACAAAGCTTCAGAGGAGGCTGCTCCCTGCGGCGCCAGCACTTCTCAGACACAACTT 1440
Qy 1441 CTTCTGCTGCTCAGAGTGGGAGGATCACTTACCCCAAGGTTCAAGACCAAA 1500
Db 1441 CTTCTGCTGCTCAGAGTGGGAGGATCACTTACCCCAAGGTTCAAGACCAAA 1500
Qy 1501 TCTTCAGAGTGGCCCTTCTGTTTCCCTGATTTGCTGTAGCTGAGGCAATGTCACAGA 1560
Db 1501 TCTTCAGAGTGGCCCTTCTGTTTCCCTGATTTGCTGTAGCTGAGGCAATGTCACAGA 1560
Qy 1561 ACCAAGAAAGCCCTCAGCGCTGAGTGTCTCCCTGACCCCTGTAATCTCTTAAGCTTAA 1620
Db 1561 ACCAAGAAAGCCCTCAGCGCTGAGTGTCTCCCTGACCCCTGTAATCTCTTAAGCTTAA 1620
Qy 1621 GATGATGAACCTTCAAAAAA 1643
Db 1621 GATGATGAACCTTCAAAAAA 1643

```

```

1      RESULT      3
2      US-08-781-891-79
3      : Sequence 79, Application US/08781891
4      : Patent No. 6090620
5      :
6      : GENERAL INFORMATION:
7      :
8      : APPLICANT: Fu, Ying-Hui
9      : APPLICANT: Yu, Chang-Eh
10     : APPLICANT: Oshima, Junko
11     : APPLICANT: Mulligan, John T.
12     : APPLICANT: Schellenberg, Gerald D.
13     : TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
14     : NUMBER OF INVENTION: WERNER'S SYNDROME
15     : NUMBER OF SEQUENCES: 209
16     :
17     : CORRESPONDENCE ADDRESS:
18     :
19     : ADDRESSEE: SEED and BERRY LLP
20     : STREET: 6300 Columbia Center, 701 Fifth Avenue
21     : CITY: Seattle
22     : STATE: Washington
23     : COUNTRY: USA
24     : ZIP: 98104-7092
25     :
26     : COMPUTER READABLE FORM:
27     :
28     : MEDIUM TYPE: Floppy disk
29     : COMPUTER: IBM PC compatible
30     : OPERATING SYSTEM: PC-DOS/MS-DOS
31     : SOFTWARE: PatentIn Release #1.0, Version #1.30
32     :
33     : CURRENT APPLICATION DATA:
34     : APPLICATION NUMBER: US/08/781,891
35     : FILING DATE: 27-DEC-1996
36     : CLASSIFICATION: 800
37     :
38     : ATTORNEY/AGENT INFORMATION:
39     : NAME: NO. 6090620tenburg Ph.D., Carol
40     : REGISTRATION NUMBER: 39,317
41     : REFERENCE/DOCKET NUMBER: 240052.419
42     :
43     : TELECOMMUNICATION INFORMATION:
44     : TELEPHONE: (206) 622-4900
45     : TELEFAX: (206) 682-6031
46     : INFORMATION FOR SEQ ID NO: 79:
47     : SEQUENCE CHARACTERISTICS:
48     : LENGTH: 87350 base pairs
49     : TYPE: nucleic acid
50     : STRANDEDNESS: single
51     : TOPOLOGY: linear
52     :
53     : US-08-781-891-79

```

CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: winpatlin (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/978,741  
 FILING DATE: 26-No. 6100076-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/792498  
 FILING DATE: 31  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Svoboda, Craig G.  
 REGISTRATION NUMBER: 39,044  
 REFERENCE/DOCKET NUMBER: P1041P1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650/225-1489  
 TELEFAX: 650/952-9881  
 INFORMATION FOR SEQ ID NO. 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5009 base pairs  
 TYPE: Nucleic Acid  
 STRANDEDNESS: Single  
 TOPOLOGY: Linear

US-08-978-741-7

	Query Match	3.8%	Score 62	DB 3	Length 5009
	Best Local Similarity	73.6%	Pred. No. 6e-07		
	Matches 109	Conservative	0	Mismatches 30	Indels 9
				Gaps 2	
QY	1	GGGAGCGTGAAGCAAGACAGACCGCTTGAAC---	GGGGGTGATGTGTGCAGTGAACCCAGA	57	
Db	4665	GGGAGCGTGAAGCAGAGAGATGGCTTGAACCTGGGAGGTGGAGGTTGCAGTGAACCCAGA	4724		
QY	58	TGTTGCCACTGCATCTCCACAGCTTGGCAACAGTCCGAGAGCGGAGACCGCTTCAAAAACA	117		
Db	4725	TGTTGCCACTGCATCTCCACAGCTTGGCAACAGTCCGAGAGCGGAGACCGCTTCAAAAACA	4778		
QY	118	AAGTCTCAAGTGGGCCAGGCTCCCTTC	145		
Db	4779	AGGAGCTTCATGTGCCCCCAATTATTC	4806		

```

Query Match          3.8%;   Score 62.8;   DB 3;   Length 87350;
Best Local Similarity 75.4%;   Pred. NO. 1.2e-06;
Matches 92;   Conservative 0;   Mismatches 27;   Indels 3;   Gaps 1.

OY      1  GGGAGCGTGAAGCAGAGAACCGCTTGAAC---GGGGGTGATGTTGCAGTGAGCCAGA 57
          |||||  ||  |||||  |||||  ||  |||||  ||  |||||  ||  |||||  ||
Db 10278 GGGAGCGTGAGGCAGAGAAATCGCTTGAACCTGGGAGCGGAGGTTGCCGTAGCCGAGA 10337
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ||

OY      58  TGTGCGCACTGCACATCCAGCCGTGGCAGAGTGCGCGCAGAGCGAGCGCCGTCTCAAAACAA 117
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
Db 10338 TTGGCGCACTGCACATCCAGCCGTGGGAGACGAGACGACACTCTGTCTCAAAAAAAAAACAA 10397
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ||

OY      118 AA 119
          ||
Db 10398 AA 10399
          .

RESULT 4
US-08-978-741-7
; Sequence 7, Application US/08978741
; Patent No. 6100076
; GENERAL INFORMATION:
; APPLICANT: Yang Wang, Michael W. Spellman
; TITLE OF INVENTION: O-Fucosyltransferase
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
;

```

```

RESULT      5
US-09-333-729A-8
: Sequence 8, Application US/09333729A
: Patent No. 6270987
: GENERAL INFORMATION:
: APPLICANT: Wang,Yang
: APPLICANT: Spellman,Michael W.
: TITLE OF INVENTION: O-Fucosyltransferase
: FILE REFERENCE: P104IPDI-Substitute
: CURRENT APPLICATION NUMBER: US/09/333,729A
: CURRENT FILING DATE: 1999-06-15
: PRIOR APPLICATION NUMBER: US 08/798,741
: PRIOR FILING DATE: 1997-11-26
: NUMBER OF SEQ ID NOS: 21
: SEQ ID NO 8
: LENGTH: 5009
: TYPE: DNA
: ORGANISM: Homo Sapien
US-09-333-729A-8

Query Match      3.8%; Score 62; DB 4; Length 5009;
Best Local Similarity 73.6%; Pred. NO. 6e-07;
Matches 109; Conservative 0; Mismatches 30; Indels 9; Gaps 2

```

```

QY 1 GGGAGGCTGACGACGAGGAGACCGCTTGAC---GGGGCTGATGTTGAGTGCAGCCACGA 57
   ||||||| ||| ||||| ||||||| ||||||| ||||||| ||||||| |||
Db 4665 gggagggctgagggcagaggaataatgcttgaacctgggaggtggaagctgagcgccaga 4724
QY 58 TGGTCCACCTGACATCCAGCCCTGGCAGACATGTCGACAGACGACGCCGTCTCAAAACNA 117
   ||||||| ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 4725 tgggtccactgcatctccagcctgggca-----acagagcgagactcctcaaaaaa 4778
QY 118 AAGTCTACAGTGGCCCAACCTCTTC 145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4779 aggaactcattggtgcccccaattttc 4806

```

```

RESULT 6
US-09-188-930-23
; Sequence 23, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 997
; TYPE: DNA
; ORGANISM: mouse
US-09-188-930-23

```

```

Query Match
Best Local Similarity 52.1%; Score 61; DB 3; Length 997;
Matches 136; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

```

```

QY 480 CCTTACCCCTTCTACTGCGACCATGATGCTTCTGTAAGATGCTGAGGACGAGGCGAC 539
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 718 ccgcagcaacaacatgctgctgtagtgatcaacaagatgctgagaagggagc 777
QY 540 CAGGACCTCTGAGGCGCTCCCGCCACCTGGTATGACTGTGCCAGCTACCGCCAT 599
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 778 caagtcactctgcygggcaacgcatcaatgtcctcaaatgcccctgagtcgcat 837
QY 600 CTACTTACTGCTATGACCACTGAAGGCTTCTGTGTGTGAGCCCTGACCTCTGA 659
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 838 caaatcactgcatatgagcagatgaagcgtctgctgtagtcatcagagagcgtgag 897
QY 660 CCTTACGACACCATGATGCTGCGCGCGCTGCGCCGCTTGGGACCGTACTGTATCAG 719
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 898 gattccagaaaggcttctggaagcgtctcttgcgcgagccattgcccagagtagcatcta 957
QY 720 CCCCCTGAGCTTATGCGGAC 740
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 958 cccaatgaggttcttgaaagac 978

```

```

RESULT 7
US-09-188-930-262
; Sequence 262, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg

```

```

; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 262
; LENGTH: 1816
; TYPE: DNA
; ORGANISM: Mouse
US-09-188-930-262

```

```

Query Match
Best Local Similarity 52.1%; Score 61; DB 3; Length 1816;
Matches 136; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

```

```

QY 480 CCTTACCCCTTCTACTGCGACCATGATGCTTCTGTAAGATGCTGAGGACGAGGCGAC 539
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 718 ccgcagcaacaacatgctgctgtagtgatcaacaagatgctgagaagggagc 777
QY 540 CAGGACCTCTGAGGCGCTCCCGCCACCTGGTATGACTGTGCCAGCTACCGCCAT 599
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 778 caagtcactctgcygggcaacgcatcaatgtcctcaaatgcccctgagtcgcat 837
QY 600 CTACTTACTGCTATGACCACTGAAGGCTTCTGTGTGTGAGCCCTGACCTCTGA 659
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 838 caaatcactgcatatgagcagatgaagcgtctgctgtagtcatcagagagcgtgag 897
QY 660 CCTTACGACACCATGATGCTGCGCGCGCTGCGCCGCTTGGGACCGTACTGTATCAG 719
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 898 gattccagaaaggcttctggaagcgtctcttgcgcgagccattgcccagagtagcatcta 957
QY 720 CCCCCTGAGCTTATGCGGAC 740
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 958 cccaatgaggttcttgaaagac 978

```

```

RESULT 8
US-09-385-982-406/c
; Sequence 406, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; EARLIER FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 406
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(591)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-406

```

```

Query Match
Best Local Similarity 80.3%; Score 60.2; DB 4; Length 591;
Matches 98; Conservative 0; Mismatches 14; Indels 10; Gaps 2;

```

RESULT 12  
PCT-US96-01314-39  
Sequence 39, Application PC/TUS9601314  
GENERAL INFORMATION:  
APPLICANT: M. Amin Arnout  
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN-  
TITLE OF INVENTION: ANTAGONISTS

NUMBER OF SEQUENCES: 78  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: Wordperfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/01314  
FILING DATE: 30-JAN-96  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/380,167  
FILING DATE: 30-JAN-95  
ATTORNEY/AGENT INFORMATION:  
NAME: John W. Freeman  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00786/267001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5137 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US96-01314-39

Query Match  
Best Local Similarity 66.9%; Score 58.6; DB 5; Length 5137;  
Matches 117; Conservative 0; Mismatches 49; Indels 9; Gaps 2;

QY 2 GGAGGCTGAAGCAGACGCGCTTGAAC---GGGGGTGATGTGACAGTGGCCCAAGT 58  
DB 4783 GGAGGTTAGGTAGAGATTTCTTGAACCTGGAGGTGGAGTTGACGTGACCAAGAT 4842  
QY 59 GGTGCACACTGCATCCAGCTTGGCAACAGTGGCAGAGCGACCGCTCTCAAAACAA 118  
DB 4843 TCGCGCATTTGACATCCAGCTGGCA-----ACACAGGAGACCTCCGTCTCAAGGAAA 4896  
QY 119 AGCTCAGAGTGGCCAGCCCTCTTCTGCGCATCCCTGGAAGTTGTGTCTTT 173  
DB 4897 AATATAAATAAAGCGGCGGCGGCGGACATCCACCCCTTGAGAGCTGTCT 4951

RESULT 13  
US-08-476-062A-39  
Sequence 39, Application US/08476062A  
Patent No. 5877375  
GENERAL INFORMATION:  
APPLICANT: Arnould, M. Amin  
TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY  
TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,062A  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/216,081  
FILING DATE: 21-MAR-1994  
APPLICATION NUMBER: 07/637,830  
FILING DATE: 04-JAN-1991  
APPLICATION NUMBER: 07/539,842  
FILING DATE: 18-JUN-1990  
APPLICATION NUMBER: 07/212,573  
FILING DATE: 28-JUN-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00786/068003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5138 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 95...3604  
US-08-476-062A-39

Query Match  
Best Local Similarity 66.9%; Score 58.6; DB 2; Length 5138;  
Matches 117; Conservative 0; Mismatches 49; Indels 9; Gaps 2;

QY 2 GGAGGCTAACAAGACGCGCTTGAAC---GGGGGTGATGTGACAGTGGCCCAAGT 58  
DB 4784 GGAGGTTAGGTAGAGATTTCTTGAACCTGGAGGTGGAGTTGACGTGACCAAGAT 4843  
QY 59 GGTGCACACTGCATCCAGCTTGGCAACAGTGGCAGAGCGGAGCGGCTCTCAAAACAA 118  
DB 4844 TCGCGCATTTGACATCCAGCTGGCA-----ACACAGGAGACTCCGTCTCAAGGAAA 4897  
QY 119 AGCTCAGAGTGGCCAGCCCTCTTCTGCGCATCCCTGGAAGTTGTGTCTTT 173  
DB 4898 AATATAAATAAAGCGGCGGCGGCGGACATCCACCCCTTGAGAGCTGTCT 4952

RESULT 14  
US-08-395-800A-7  
Sequence 7, Application US/08395800A  
Patent No. 5807732  
GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B  
TITLE OF INVENTION: SAME  
TITLE OF INVENTION: METHOD FOR PRODUCING THE SAME AND A METHOD OF  
TITLE OF INVENTION: GENOTYPING A PERSON  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: USA

ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/395,800A  
FILING DATE: 28-FEB-1995  
CLASSIFICATION: 435  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2115 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 64..1092  
US-08-395-800A-7

Query Match 3.6%; Score 58.4; DB 1; Length 2115;  
Best Local Similarity 79.5%; Pred. No. 3.8e-06;  
Matches 97; Conservative 0; Mismatches 16; Indels 9; Gaps 2;

QY 1 GGGAGCTGACGACGACGCTGAC--GGGGGTGATGTTGACGTAGCCAGA 57  
DB 1422 GGGAGGCTGAGACGATGACCTTGACCCAGAGGCGGAGGTGACGACCA 1481  
QY 58 TGGTGGCACTGACCTGCGACAGTCGCGAGAGCGCGCTCTCAAAACA 117  
DB 1482 TGGTGGCGGTGACCTGCGACAGTCGCGAGAGCGCGCTCTCAAAACA 1535  
QY 118 AA 119  
DB 1536 AA 1537

RESULT 15  
US-08-310-356-36/C  
Sequence 36, Application US/08310356  
Patent No. 5648243  
GENERAL INFORMATION:  
APPLICANT: Hurwitz, David R  
APPLICANT: Nathan, Margret  
APPLICANT: Shaul, Moshe  
TITLE OF INVENTION: Transgenic Protein Production  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Legal Department  
STREET: 500 Arcola Road  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh System 7.0  
SOFTWARE: Microsoft Word Version 5.0 (Patentln)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/310,356  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/737,853  
FILING DATE: 31-JUL-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Goodman, Rosanne  
REGISTRATION NUMBER: 32,534  
REFERENCE/DOCKET NUMBER: A0856  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 454-3817  
TELEFAX: (215) 454-3808  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19011 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: exon  
LOCATION: 1737..1854  
OTHER INFORMATION: /number=1  
OTHER INFORMATION: /citation=(11)  
FEATURE:  
NAME/KEY: intron  
LOCATION: 1855..2563  
OTHER INFORMATION: /label=intron1  
OTHER INFORMATION: /citation=(11)  
FEATURE:  
NAME/KEY: exon  
LOCATION: 2564..2621  
OTHER INFORMATION: /number=2  
OTHER INFORMATION: /citation=(11)  
FEATURE:  
NAME/KEY: intron  
LOCATION: 2622..4075  
OTHER INFORMATION: /number=2  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4076..4208  
OTHER INFORMATION: /number=3  
OTHER INFORMATION: /citation=(11)  
FEATURE:  
NAME/KEY: intron  
LOCATION: 4209..6040  
OTHER INFORMATION: /number=3  
FEATURE:  
NAME/KEY: exon  
LOCATION: 6041..6252  
OTHER INFORMATION: /number=4  
FEATURE:  
NAME/KEY: intron  
LOCATION: 6253..6801  
OTHER INFORMATION: /number=4  
FEATURE:  
NAME/KEY: exon  
LOCATION: 6802..6934  
OTHER INFORMATION: /number=5  
FEATURE:  
NAME/KEY: intron  
LOCATION: 6935..7758  
OTHER INFORMATION: /number=5  
FEATURE:  
NAME/KEY: exon  
LOCATION: 7759..7856  
OTHER INFORMATION: /number=6  
FEATURE:  
NAME/KEY: intron  
LOCATION: 7857..9443  
OTHER INFORMATION: /number=6  
FEATURE:  
NAME/KEY: exon  
LOCATION: 9444..9573  
OTHER INFORMATION: /number=7  
FEATURE:  
NAME/KEY: intron

```
LOCATION: 9574..10866
FEATURE:
OTHER INFORMATION: /number- 7
NAME/KEY: exon
LOCATION: 10867..11081
OTHER INFORMATION: /number- 8
FEATURE:
NAME/KEY: intron
LOCATION: 11082..12480
OTHER INFORMATION: /number- 8
FEATURE:
NAME/KEY: exon
LOCATION: 12481..12613
OTHER INFORMATION: /number- 9
FEATURE:
NAME/KEY: intron
LOCATION: 12614..13701
OTHER INFORMATION: /number- 9
FEATURE:
NAME/KEY: exon
LOCATION: 13702..13799
OTHER INFORMATION: /number- 10
FEATURE:
NAME/KEY: intron
LOCATION: 13800..14976
OTHER INFORMATION: /number- 10
FEATURE:
NAME/KEY: exon
LOCATION: 14977..15115
OTHER INFORMATION: /number- 11
FEATURE:
NAME/KEY: intron
LOCATION: 15116..15533
OTHER INFORMATION: /number- 11
FEATURE:
NAME/KEY: exon
LOCATION: 15534..15757
OTHER INFORMATION: /number- 12
FEATURE:
NAME/KEY: intron
LOCATION: 15758..16949
OTHER INFORMATION: /number- 12
FEATURE:
NAME/KEY: exon
LOCATION: 16950..17082
OTHER INFORMATION: /number- 13
FEATURE:
NAME/KEY: intron
LOCATION: 17083..17696
OTHER INFORMATION: /number- 13
FEATURE:
NAME/KEY: exon
LOCATION: 17697..17764
OTHER INFORMATION: /number- 14
FEATURE:
NAME/KEY: intron
LOCATION: 17765..18534
OTHER INFORMATION: /number- 14
FEATURE:
NAME/KEY: CDS
LOCATION: 1041(1776..1854, 2564..2621, 4076..4208,
6041..6252, 6802..6934, 7759..7856, 9444..9573,
10867..11081, 12481..12613, 13702..13799,
14977..15115, 15534..15757, 16950..17082,
17697..17741)
OTHER INFORMATION: /product= "human serum albumin"
OTHER INFORMATION: /citation= ([1])
FEATURE:
NAME/KEY: exon
LOCATION: 18535..18697
OTHER INFORMATION: /number- 15
FEATURE:
NAME/KEY: 3'UTR
```

```
LOCATION: 17742..18697
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1737..1775
PUBLICATION INFORMATION:
AUTHORS: Minghetti, P P
AUTHORS: Ruffner, D E
AUTHORS: Kuang, W-J
AUTHORS: Dennison, O E
AUTHORS: Hawkins, J W
AUTHORS: Beatlie, W G
AUTHORS: Dugalczyk, A
TITLE: MOLECULAR STRUCTURE OF THE HUMAN ALBUMIN
TITLE: GENE IS REVEALED BY NUCLEOTIDE SEQUENCE WITHIN
TITLE: 911-22 OF CHROMOSOME 4
JOURNAL: J. Biol. Chem.
PAGES: 6747-6757
VOLUME: 261
DATE: 1986
RELEVANT RESIDUES IN SEQ ID NO: 36: FROM 1 TO 19011
US-08-310-356-36
```

```
Query Match          3.6% Score 58.4; DB 1; Length 19011;
Best Local Similarity 79.5% Pred. No. 9.4e-06;
Matches 97; Conservative 0; Mismatches 16; Indels 9; Gaps 2;

QY 1 GGGAGCGCTGAGCAGAGAACCGCTTGAAC--GGGGGTGATGTTGCAGTGAGCCAGA 57
    ||||||| || ||||| ||||| || ||||| ||||| ||||| ||||| |||||
DB 15419 GGGAGCGCTGAGCAGAGAAATGCTTGAACCTGGAGGGGAGTGGCAGTGAGCCAGA 15360
QY 58 TGGTGCACTGCACCTCCAGCCCTGGCAGACAGTCCGCGAGAGCCGCTCTCAAAACAA 117
    || || ||||||| ||||||| || ||||| || ||||| || ||||| || |||||
DB 15359 TGGCACCACTGCACCTCCAGCCCTGGACA-----ACAAAGCGAGACTCCGCTCAAAAAA 15306
QY 118 AA 119
    ||
DB 15305 AA 15304
```

Search completed: April 19, 2002, 21:44:45  
Job time: 7344 sec

Sat Apr 20 10:03:16 2002

us-09-840-787-68.rni

---

Page 10

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2002, 19:38:36 ; Search time 1474.2 Seconds  
(Without alignments)  
11976.200 Million cell updates/sec

Title: US-09-840-787-68

Perfect score: 1643  
Sequence: 1 GGGAGCGCTGAAGCAGAGAA.....GATGAACCTTCAAAAAAAAA 1643

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hic:\*  
10: gb\_est1:\*  
11: gb\_est2:\*  
12: gb\_hic:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rod:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1388.8	84.5	1402	12	AF119864 Homo sapi
2	908.8	55.3	963	10	AL519986 AL519986
3	864	52.6	910	10	AL570513 AL570513
4	838.8	51.1	884	10	AL572559 AL572559
5	824.4	50.2	1001	10	AL519987 AL519987
6	823	50.1	921	10	AL529906 AL529906
7	814	49.5	925	10	AL516722 AL516722
8	812.8	49.5	911	10	AL528266 AL528266
9	806.6	49.1	1450	12	AK007934 Mus muscu
10	798.8	48.6	816	11	BG470240 602533758
11	796.4	48.5	885	10	AL514846 AL514846
12	795.8	48.4	940	10	AL523298 AL523298

C 13	790.4	48.1	829	10	AL563104
C 14	789.4	48.0	1371	12	AK007276
C 15	786.6	47.9	795	10	AL571090
C 16	786.2	47.9	862	11	BG470248
C 17	774.4	47.1	901	10	AL545150
C 18	771.4	47.0	891	10	AL522522
C 19	770.2	46.9	1453	12	AK002381
C 20	769.4	46.8	792	11	BG470228
C 21	765.2	46.6	880	10	AL515130
C 22	764	46.5	937	10	AL578185
C 23	763	46.4	917	10	AL544360
C 24	758.6	46.2	1491	12	AK019396
C 25	757.2	46.1	791	10	AL522521
C 26	755	46.0	762	10	AL522849
C 27	753.6	45.9	848	10	BE740464
C 28	750.8	45.7	911	11	BG337246
C 29	743.2	45.2	945	10	AL523299
C 30	739	45.0	783	10	AL529905
C 31	739	44.9	774	11	BG831987
C 32	737.8	44.9	841	11	BI259744
C 33	735.8	44.8	781	10	AI005021
C 34	735.4	44.7	847	11	BI333101
C 35	734.6	44.7	918	11	BG574617
C 36	729.2	44.4	775	10	AL515129
C 37	728.6	44.3	834	10	AL514790
C 38	728.4	44.3	802	11	BG775712
C 39	724.4	44.1	810	10	AI279886
C 40	723.8	44.1	810	11	BG749617
C 41	723.2	44.0	881	11	BI085334
C 42	710.6	43.3	809	11	BI092202
C 43	710.2	43.2	814	10	AA688509
C 44	708.8	43.1	814	10	AL516121
C 45	706.2	43.0	731	10	AL516121

## ALIGNMENTS

RESULT 1  
AF119864 1402 bp mRNA HTC 08-MAY-2001  
LOCUS AF119864  
DEFINITION Homo sapiens PRO2163 mRNA, complete cds.  
ACCESSION AF119864  
VERSION AF119864.1 GI:7770164  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M. and He,F.  
TITLE Functional prediction of the coding sequences of 79 new genes deduced by analysis of cDNA clones from human fetal liver  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1402)

TITLE Direct Submission  
JOURNAL Submitted (13-JAN-1999) Department of Experimental Hematology, Institute of Radiation Medicine, Beijing Taping Road 27, Beijing, Beijing 100850, P. R. China  
AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M. and He,F.  
FEATURES  
source location/Qualifiers  
1..1402  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="FLB8041"  
/tissue\_type="liver"  
/dev\_stage="fetus"  
4..1017  
/codon\_start=1  
/product="PRO2163"  
/protein\_id="AAF69618.1"

```

/db_xref="GI:7770165"
/translation="MVASGTGAVTSLFMTPLDVVKVRLQSORPMASLELPSRLMS
LSTYKMKCLILYNGVLEPLTICPNAGARCAMPDOPREFGTGMAFVIVHEGRRLM
SGLPARTLVMTVPNAIFYTAYDOLAKILCRALITSDIAXMVNAGALRLITVYISPL
ELMKRLQAOHVSTRELGACVKTAVAGGSRSLMGLPALKRVPFSAITWETELV
KSMNLGRPRDQTSVSGSVFVAGISRTVAVALPLPEFVVTQORVALGAEAVVNL
HVDSTWLLRRIRAESGCKGLFAGFLPRITIKAAFPSCAIMISTEEFKSPFORLMDRL
LGG"
BASE COUNT      261 a      432 c      401 g      308 t
ORIGIN
Query Match      84.5%; Score 1388.8; DB 12; Length 1402;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1393; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 244 CAAATGTGGGCTCAGGACCGGGGCTGTGTACCTCTCTCTCATGACACCCCTGGAC 303
D 1 CAAATGTGGGCTCAGGACCGGGGCTGTGTACCTCTCTCTCATGACACCCCTGGAC 60
QY 304 GTGGTGAAGGTTGGCTGACGTCAGCGGCTCCATGGCAGGAGACTGATGGCTTCC 363
D 61 GTGGTGAAGGTTGGCTGACGTCAGCGGCTCCATGGCAGGAGACTGATGGCTTCC 120
QY 364 TCAGACTGTGGAGCTCTCTCTATACCAATGGAAGTGCCTCTGTATTGCAATGTGTC 423
D 121 TCAGACTGTGGAGCTCTCTCTATACCAATGGAAGTGCCTCTGTATTGCAATGTGTC 180
QY 424 CTGAGGCTGTACCTGTGCGCAATGAGTGGCCGCTGTGGCAGCTGTTCAAGACCT 483
D 181 CTGAGGCTGTACCTGTGCGCAATGAGTGGCCGCTGTGGCAGCTGTTCAAGACCT 240
QY 484 ACCGCTTCACTGACGACCATGATGCTCTTCTGAAGATCTGAGGACAGAGGACACAG 543
D 241 ACCGCTTCACTGACGACCATGATGCTCTTCTGAAGATCTGAGGACAGAGGACACAG 300
QY 544 ACCCTCTGAGAGCGGCTCCCGGCCACCTGTGTATGACTGTGCCAGCTACCCCATCTAC 603
D 301 ACCCTCTGAGAGCGGCTCCCGGCCACCTGTGTATGACTGTGCCAGCTACCCCATCTAC 360
QY 604 TTCACTGCTTATGACCAATGAGGCTTCTGTGTGTGTCAGGCTTACCTGTGACCTC 663
D 361 TTCACTGCTTATGACCAATGAGGCTTCTGTGTGTGTCAGGCTTACCTGTGACCTC 420
QY 664 TAGCAGCCATGTGTGCTGCGCGCTGGCCGCTTGGGACCGTGTATGTATGACGCC 723
D 421 TAGCAGCCATGTGTGCTGCGCGCTGGCCGCTTGGGACCGTGTATGTATGACGCC 480
QY 724 CTGAGGCTTATGCGGACAAAGCTGACGAGCTGAGATGTGTACCGGAGAGCTGGGTGCC 783
D 481 CTGAGGCTTATGCGGACAAAGCTGACGAGCTGAGATGTGTACCGGAGAGCTGGGTGCC 540
QY 784 TGTGTTCGACATGAGTGGCTCAGGAGTGGCTGAGGCTGACTGTGGCTGGGAGGCC 843
D 541 TGTGTTCGACATGAGTGGCTCAGGAGTGGCTGAGGCTGACTGTGGCTGGGAGGCC 600
QY 844 ACTGCCCTTCAGATGTGCCCTTCTCAGCCCTGTACTGTGTTCACTATGAGCTGTGAAG 903
D 601 ACTGCCCTTCAGATGTGCCCTTCTCAGCCCTGTACTGTGTTCACTATGAGCTGTGAAG 660
QY 904 AGTGGCTTCATGAGGCTCAGGCGCAAGAGACGACTTCTGTGGGATAGCTTGTGGCT 963
D 661 AGTGGCTTCATGAGGCTCAGGCGCAAGAGACGACTTCTGTGGGATAGCTTGTGGCT 720
QY 964 GGTGGCATCTCAGGAGCGGTGGCTGACGTGACTTACCTTTGACGTGGTAAAGAC 1023
D 721 GGTGGCATCTCAGGAGCGGTGGCTGACGTGACTTACCTTTGACGTGGTAAAGAC 780
QY 1024 CAAAGCCAGGTGCTCTGAGGAGCATGAGGCTGTGAGAGTGAACCCCTGATGTGAC 1083
D 781 CAAAGCCAGGTGCTCTGAGGAGCATGAGGCTGTGAGAGTGAACCCCTGATGTGAC 840
QY 1084 TCACACTGTGCTGTGCTGAGAGATCCGGGCCGAGTGGGACACCAAGGACTTTTGCA 1143

```

```

D 841 TCACACTGTGCTGTGCTGAGAGATCCGGGCCGAGTGGGACACCAAGGACTTTTGCA 900
QY 1144 GGCCTTCCTCTCGATATATGAAGGCTGCCCTCTCTGCTGATCATGATAGACCTAT 1203
D 901 GGCCTTCCTCTCGATATATGAAGGCTGCCCTCTCTGCTGATCATGATAGACCTAT 960
QY 1204 GAGTTCGCAAAAGCTTCTTCAGAGGCTGAACAGGACCGGCTTCTGGGCGGTGAAG 1263
D 961 GAGTTCGCAAAAGCTTCTTCAGAGGCTGAACAGGACCGGCTTCTGGGCGGTGAAG 1020
QY 1264 GGGCAGAGGACGACAGGACCCGCTCTCTCCACGATGGGAGAGAGGACAGACCA 1323
D 1021 GGGCAGAGGACGACAGGACCCGCTCTCTCCACGATGGGAGAGAGGACAGACCA 1080
QY 1324 GCGAAGTGCCTTCTTCAGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1383
D 1081 GCGAAGTGCCTTCTTCAGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
QY 1384 AAGCTCCAGGACAGGAGCTGTCCCTCTGGGCGGCCAGCACTTCTTCAGACAACTTCT 1443
D 1141 AAGCTCCAGGACAGGAGCTGTCCCTCTGGGCGGCCAGCACTTCTTCAGACAACTTCT 1200
QY 1444 CCTGCTGCTCCAGTGTGGGATATCACTTACCAACCCCAAGTTAAACCAACTCT 1503
D 1201 CCTGCTGCTCCAGTGTGGGATATCACTTACCAACCCCAAGTTAAACCAACTCT 1260
QY 1504 TCAGACTGCCCCCTGCTGCTTCCCTGTGTGCTGTGAGTGGGAGCTGTCCAGAAC 1563
D 1261 TCAGACTGCCCCCTGCTGCTTCCCTGTGTGCTGTGAGTGGGAGTGTCCAGAAC 1320
QY 1564 AAGAAGCCCTAGCCTGTGTGTGCTGCTGCTGACCTGTGTAACTTAACTTAAAGAT 1623
D 1321 AAGAAGCCCTAGCCTGTGTGTGCTGCTGCTGACCTGTGTAACTTAACTTAAAGAT 1380
QY 1624 GATGACTTCAAAAAA 1643
D 1381 GATGAAAAA 1400

```

```

RESULT 2
AL519986/c 963 bp mRNA EST 13-FEB-2001
LOCUS AL519986 LTL_NFL004_NBC2 Homo sapiens cDNA clone CSDB005YF09 3
DEFINITION prime, mRNA sequence.
ACCESSION AL519986
VERSION AL519986.1 GI:12783479
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 963)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.

```

## FEATURES

```

source
1. 963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDB005YF09"
/clone_id="LTL_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain. Vector: pCMVSPORT 6. 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6

```

vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiang@lifetech.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 206 a 288 c 298 g 168 t 3 others  
ORIGIN

Query Match 55.38; Score 908.8; DB 10; Length 963;  
Best Local Similarity 98.28; Pred. No. 4.5e-209;  
Matches 947; Conservative 3; Mismatches 11; Indels 3; Gaps 3;

QY 650 TGACCTCTGACCTTACGACCCATGAGGCTGCGCGCTGGCCGCTTGGGACCGTGA 709  
|||||  
Db 963 TGACCTCTGACCTTACGACCCATGAGG-TKCGCGCTGGCCGCTTGGGACCGTGA 905  
QY 710 CTGTGATCAGCCCGCTGAGCTTATGCGGACAAAGCTGCAGGCTCAGCATGTGCTAAC 769  
|||||  
Db 904 CTGTGATCAGCCCGCTGAGCTTATGCGGACAAAGCTGCAGGCTCAGCATGTGCTAAC 845  
QY 770 GGGAGCTGGGCTGCTGTGTTGAACTGCAGTGGCTCAGGCTGGCGCTCAGTGTGGC 829  
|||||  
Db 844 GGGAGCTGGGCTGCTGTGTTGAACTGCAGTGGCTCAGGCTGGCGCTCAGTGTGGC 785  
QY 830 TGGGCTGGGGCCCGCTGCTGAGATGTGCTTCTGAGCCCTGACTGATGTTCAACT 889  
|||||  
Db 784 TGGGCT-GGGCGCCACCTGCTTCTGAGATGTGCTTCTGAGCCCTGACTGATGTTCAACT 726  
QY 890 ATGAGCTGTGTAAAGCTGTGCTCAATGGGCTCAGGCGGCAAGACAGACTTCTGTGGCA 949  
|||||  
Db 725 ATGAGCTGTGTAAAGCTGTGCTCAATGGGCTCAGGCGGCAAGACAGACTTCTGTGGCA 666  
QY 950 TGACCTTGTGGCTGTGTGGCTATCTCAGGAGCGTGGCTGCAGTGTGACTTACCCTTTG 1009  
|||||  
Db 665 TGACCTTGTGGCTGTGTGGCTATCTCAGGAGCGTGGCTGCAGTGTGACTTACCCTTTG 606  
QY 1010 ACGTGTGTAAAGACCAAGCGGCTGTGGGAGCGATGAGGCTGTGAGAGTGAAC 1069  
|||||  
Db 605 ACGTGTGTAAAGACCAAGCGGCTGTGGGAGCGATGAGGCTGTGAGAGTGAAC 546  
QY 1070 CCGTGTGTGTGACTTCCAGCTGTGCTGTGCGGAGATCCGGGCGGAGTGGGACCA 1129  
|||||  
Db 545 CCGTGTGTGTGACTTCCAGCTGTGCTGTGCGGAGATCCGGGCGGAGTGGGACCA 486  
QY 1130 AGGAGCTTTTGCAGGCTTCTTCTCGGATCATCAAGGCTGCCCTCTGTGCTATCA 1189  
|||||  
Db 485 AGGAGCTTTTGCAGGCTTCTTCTCGGATCATCAAGGCTGCCCTCTGTGCTATCA 426  
QY 1190 TGATCAGACCTATGAGTGTGGCAAAAGCTTCTCCAGAGGCTGAACCGGAGGCTTC 1249  
|||||  
Db 425 TGATCAGACCTATGAGTGTGGCAAAAGCTTCTCCAGAGGCTGAACCGGAGGCTTC 366  
QY 1250 TGGGCGGCTGAAAGGGGCAAGAGGCAAGCCGCTCTCCACAGGATGGGAGAGG 1309  
|||||  
Db 365 TGGGCGGCTGAAAGGGGCAAGAGGCAAGCCGCTCTCCACAGGATGGGAGAGG 306  
QY 1310 CAGAGAGAGACCCAGCAAGTGTCTTCTCAGACACTGAGGAGAGGGGCTTGTTCCT 1369  
|||||  
Db 305 CAGAGAGAGACCCAGCAAGTGTCTTCTCAGACACTGAGGAGAGGGGCTTGTTCCT 246  
QY 1370 TCCCTCCGCGGAGCAAGTGTCCAGGAGGAGGCTGTCCCTGCGGAGGCGGACACTTCTC 1429  
|||||  
Db 245 TCCCTCCGCGGAGCAAGTGTCCAGGAGGAGGCTGTCCCTGCGGAGGCGGACACTTCTC 186  
QY 1430 AGACACAATCTTCTGCTGCTCCAGTCTGTGGGAGTATCACTTACCACCCCAAGT 1489  
|||||  
Db 185 AGACACAATCTTCTGCTGCTCCAGTCTGTGGGAGTATCACTTACCACCCCAAGT 126  
QY 1490 TCAAGACCAATCTTCCAGCTGCGCCCTTGTCTTCTGTCTGTGTGAGTGGGCA 1549  
|||||  
Db 125 TCAAGACCAATCTTCCAGCTGCGCCCTTGTCTTCTGTGTGAGTGGGCA 67

QY 1550 TGTCTCAGAGCAACCAAGAGCCCTCAGCCTGTGCTAGTCTCCCTGACCTTGTAAATTC 1609  
|||||  
Db 66 TGTCTCAGAGCAACCAAGAGCCCTCAGCCTGTGCTAGTCTCCCTGACCTTGTAAATTC 7  
QY 1610 TTAAT 1613  
|||||  
Db 6 TTAAT 3

## RESULT 3

AL570513/c 910 bp mRNA EST 16-FEB-2001  
AL570513 LTI\_NF1006.PL2 Homo sapiens cDNA clone CS0D1018YN07 3  
DEFINITION prime, mRNA sequence.

ACCESSION AL570513  
VERSION AL570513.1 GI:12926896

## KEYWORDS

EST.

## SOURCE

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 910)

Li, W.B., Gruber, C., Jesses, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: [secref@genoscope.cns.fr](mailto:secref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

## FEATURES

source

1..910

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0D1018YN07"

/clone\_id="LTI\_NF1006.PL2"

/tissue\_type="placenta"

/note="Vector: PCMVSPORT 6; Site\_1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the PCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Feng Liang Life Technologies,

a division of invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : [filiang@lifetech.com](mailto:filiang@lifetech.com) URL : <http://fulllength.invitrogen.com>

BASE COUNT 197 a 260 c 272 g 156 t 25 others  
ORIGIN

Query Match 52.68; Score 864; DB 10; Length 910;  
Best Local Similarity 96.08; Pred. No. 2.9e-198;  
Matches 868; Conservative 24; Mismatches 11; Indels 1; Gaps 1;  
QY 626 AGGCTTCTGTGTGTGAGCCCTGACCTTGAACCTTACGACCCATGAGTGGTGGCG 685  
|||||  
Db 903 ARGCTTCTGTGTGTGAGCCCTGACCTTGAACCTTACGACCCATGAGTGGTGGCG 844  
QY 686 CGGTGGCCGCTTGGGACCGCTGACTGTGATCAGCCCTGAGCTTATGGGACAAAGC 745  
|||||  
Db 843 CGGTGGCCGCTTGGGACCGCTGACTGTGATCAGCCCTGAGCTTATGGGACAAAGC 784  
QY 746 TGCAGGCTCAGACATGTGTGTACCGGAGGCTGTGTGTGAGTGTGAGTGTGCTC 805  
|||||  
Db 783 TGCAGGCTCAGACATGTGTGTACCGGAGGCTGTGTGTGAGTGTGAGTGTGCTC 724  
QY 806 AGGT 865  
|||||  
Db 723 AGGT 665  
QY 866 TGTGAGCCCTGTACTGTGTTCACTATGAGCTGTGTGAAGAGTGTGTGTGTGTGTGTGT 925  
|||||

```

Db 664 TCTCAGCCCTGACTGCTTCACTATGAGCTGTGTAAGMGTGCTCAATGGGCTCAGGC 605
QY 926 CGAAGGACACACACTTCTGTGGGAGTGGAGCTTGTGGGCTGTGGGATCTCAGGAGAGGTGG 985
Db 604 CGAAGGACACACACTTCTGTGGGAGTGGAGCTTGTGGGCTGTGGGATCTCAGGAGAGGTGG 545
QY 986 CTGCAAGTCTGACTCTTACCTCTTGGAGTGTGTAAGAGACCAAGGAGGCTGTGGGAG 1045
Db 544 CTGCAAGTCTGACTCTTACCTCTTGGAGTGTGTAAGAGACCAAGGAGGCTGTGGGAG 485
QY 1046 CGATGGAGGCTGTGAGATGAACCCCTGCAATGTGAGTCCACCTGCTGCTGCGGA 1105
Db 484 CGATGGAGGCTGTGAGATGAACCCCTGCAATGTGAGTCCACCTGCTGCTGCGGA 425
QY 1106 GGATCCGGGCGGAGTCCGGGACCAAGGAGCTCTTGGAGGCTCTTCTGATCATCA 1165
Db 424 GGATCCGGGCGGAGTCCGGGACCAAGGAGCTCTTGGAGGCTCTTCTGATCATCA 365
QY 1166 AGGCTGCCCTCTGCTGCTGATCATGATCAGCACCTATGAGTTCGGCAAAAGCTTCTTC 1225
Db 364 AGGCTGCCCTCTGCTGCTGATCATGATCAGCACCTATGAGTTCGGCAAAAGCTTCTTC 305
QY 1226 AGAGGCTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1285
Db 304 AGAGGCTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 245
QY 1286 TCTCTCCACGAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1345
Db 244 TCTCTCCACGAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 185
QY 1346 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1405
Db 184 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 125
QY 1406 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1465
Db 124 TTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 65
QY 1466 TCATCACTTACCAACCCCAAGTTCAGAGCAACCAATCTTCAGAGTCCCTGCTGTTT 1525
Db 64 TCATCACTTACCAACCCCAAGTTCAGAGCAACCAATCTTCAGAGTCCCTGCTGTTT 5
QY 1526 CCTT 1529
Db 4 CCTT 1

```

RESULT 4  
 AL572559/c 884 bp mRNA EST 16-FEB-2001  
 LOCUS AL572559 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CS0D108YD05 3  
 DEFINITION prime, mRNA sequence.  
 ACCESSION AL572559  
 VERSION AL572559.1 GI:12930946  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 884)  
 AUTHORS Li, M. B., Gruber, C., Jessee, J. and Polayres, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
 Location/Qualifiers  
 1..884  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0D108YD05"  
 /clone\_id="LTI\_NFL006\_PL2"

```

/life_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com url :
http://fulllength.invitrogen.com"
BASE COUNT 194 a 262 c 268 g 155 t 5 others
ORIGIN
Query Match 51.1%; Score 838.8; DB 10; Length 884;
Best Local Similarity 99.28; Pred. No. 3.6e-192;
Matches 859; Conservative 4; Mismatches 1; Indels 2; Gaps 2;
715 ATCAGCCCTGAGACTTATGCGGCAAAAGCTGAGGCTCAGCATGTGTGTCAGGGAG 774
Db 884 ATCAGCCCTGAGACTTATGCGGCAAAAGCTGAGGCTCAGCATGTGTGTCAGGGAG 825
QY 775 CTGGGCTGCTGTTCGAACTGCAATGGCTCAGAGGTGGCTGGCTCACTGTGGCTGGGC 834
Db 824 CTGGGCTGCTGTTCGAACTGCAATGGCTCAGAGGTGGCTGGCTCACTGTGGCTGGGC 765
QY 835 TGGGCGCCCACTGCGCTTCGAGATGTGCTTCAGGCTGTGATGTTCAACTATGAG 894
Db 764 T-TGGGCCCACTGCGCTTCGAGATGTGCTTCAGGCTGTGATGTTCAACTATGAG 706
QY 895 CTGTGAAAGAGTGGCTCAATGGGCTCAGGCCAAGAACCAAGCTTGTGTGGGATGAGC 954
Db 705 CTGTGAAAGAGTGGCTCAAT-TGCTCAGGCCAAGAACCAAGCTTGTGTGGGATGAGC 647
QY 955 TTTGTGGTGTGGATGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1014
Db 646 TTTGTGGTGTGGATGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 587
QY 1015 GTAAAGACCCCAAGGCGAGGTGCTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1074
Db 586 GTAAAGACCCCAAGGCGAGGTGCTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 527
QY 1075 CATGTGACTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1134
Db 526 CATGTGACTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 467
QY 1135 CTCTTGAGAGGCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1194
Db 466 CTCTTGAGAGGCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 407
QY 1195 AGCACCTATGAGTTCGGCAAAAGCTTTCAGAGGCTGAAACAGAGACCGGCTTGTGGGC 1254
Db 406 AGCACCTATGAGTTCGGCAAAAGCTTTCAGAGGCTGAAACAGAGACCGGCTTGTGGGC 347
QY 1255 GGCTGAAGGGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1314
Db 346 GGCTGAAGGGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 287
QY 1315 GGAGAGCCAGCAAGTGGCTTTCCTCAGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1374
Db 286 GGAGAGCCAGCAAGTGGCTTTCCTCAGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 227
QY 1375 CCGGCGCAAGAGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1434
Db 226 CCGGCGCAAGAGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 167
QY 1435 CAACCTTCTCTGCTGCTCAGTGTGGGATCATCACTTACCCACCCCAAGTTCAG 1494
Db 166 CAACCTTCTCTGCTGCTCAGTGTGGGATCATCACTTACCCACCCCAAGTTCAG 107
QY 1495 ACCAAATCTTCAGAGTGGCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1554

```

Db	106	ACCAATATCTTCAGCTGCCCCCTTCGTGTTTCCATGTGTTTCTGTAGCTGGGCAATGTT	47
0y	1555	CCAGGACCAAGAACGCTCAGCCG 1580	
Db	46	CCAGGACCAAGAACGCTCAGCCG 21	
RESULT	5		
AL519987			
LOCUS	AL519987	1001 bp	mRNA
DEFINITION	AL519987	L1.NFL004_NBC2	Homo sapiens cDNA clone CS0DB005YP09 5
ACCESSION	AL519987		
VERSION	AL519987.1	GI:12783480	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Enxarxota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1. (bases 1 to 1001)		
JOURNAL	L1.N.B., Gruber.C., Jessse.J., and Polayes.D.		
COMMENT	Full-length cDNA libraries and normalization		
	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	Bp 191 91006 EVRY cedex - France		
	Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
FEATURES			
source	1..1001		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0DB005YP09"		
	/clone_11b="L1.NFL004_NBC2"		
	/sex="male"		
	/issue_type="neuroblastoma cells"		
	/lab_host="DH10B"		
	/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA		
	was primed with a NotI-oligo(dt) primer. Five prime end		
	enriched, double-stranded cDNA was digested with Not I and		
	cloned into the Not I and Eco RV sites of the pCMVSPORT 6		
	vector. Library was normalized. Library was constructed		
	by Life Technologies. Contact : Feng liang life		
	Technologies, a division of Invitrogen 9800 Medical Center		
	Drive Rockville, Maryland 20850, USA Fax : (1) 301 610		
	8371 Email : fliang@lifestech.com URL :		
	http://fulllength.invitrogen.com"		
BASE COUNT	167 a 309 c 315 g 208 t		2 others
ORIGIN			
Query Match	50.2%	Score 824.4;	DB 10; Length 1001;
Best Local Similarity	98.7%	Pred. No. 1,1e-188;	
Matches 894;	Conservative 0;	Mismatches 6;	Indels 6; Gaps 6;
0y	187	CAGGACTGACCTCAATGATGGCTACACGACCCCTGGGGATCAGGCCCTCCAGCAA	246
Db	83	CAGGACTGACCTCAATGATGGCTACACGACCCCTGGGGATCAGGCCCTCCAGCAA	142
0y	247	ATGCTGACCTCAGGACCGGGGCTGTGTACCTCTCTTCATGACACCCCTGGACGTG	306
Db	143	ATGCTGACCTCAGGACCGGGGCTGTGTACCTCTCTTCATGACACCCCTGGACGTG	202
0y	307	GTAAGGTTGCGCTGACCTCAGGACCGGGGCTGTGTACCTCTCTTCATGACACCCCTGGACGTG	366
Db	203	GTAAGGTTGCGCTGACCTCAGGACCGGGGCTGTGTACCTCTCTTCATGACACCCCTGGACGTG	262
0y	367	AGACTGTGACCTCTCTTATACCAATGGAAGTGCCTCTGTATTCGAATGTGTCTTG	426
Db	263	AGACTGTGACCTCTCTTATACCAATGGAAGTGCCTCTGTATTCGAATGTGTCTTG	322
0y	427	GAGCTCTGTACCTGTGCGCAATGAGTGCCTGTGCTGACCTGGTTTCAGACCTTACC	486
Db	323	GAGCTCTGTACCTGTGCGCAATGAGTGCCTGTGCTGACCTGGTTTCAGACCTTACC	382

OY	487	CGCTTCACATGGACCATGGATGGATGCCCTTCGTGAAGATCGAGAGCAACGAGGGACACGAGAC	546
Db	383	CGCTTCACATGGACCATGGATGGATGCCCTTCGTGAAGATCGAGAGCAACGAGGGACACGAGAC	442
OY	547	CTCTGGAGCGGCGCTCCCGCCACCCCTGGTGTGATGTGACTGTGCCAGGTACCCGCATCTACTTC	606
Db	443	CTCTGGAGCGGCGCTCCCGCCACCCCTGGTGTGATGTGACTGTGCCAGGTACCCGCATCTACTTC	502
OY	607	ACTGCTTATGACCAACTAAAGGCTTCTCTGTGTGTGTGAGCGCCCTGACCTCTGACCTCTAC	666
Db	503	ACTGCTTATGACCAACTAAAGGCTTCTCTGTGTGTGTGAGCGCCCTGACCTCTGACCTCTAC	562
OY	667	GCACCCATGTGTGGCTGGGCGCGTGGCCCGCTTGGGGCACCGTGACGTGTGATGACGCCCTG	726
Db	563	GCACCCATGTGTGGCTGGGCGCGTGGCCCGCTTGGGGCACCGTGACGTGTGATGACGCCCTG	621
OY	727	GAGCTTATGCGGACAAAGCTCGAGGCTTCAGCATGTGTCTACCGGGAGCTGGGTGCCGT	786
Db	622	GAGCTTATGCGGACAAAGCTCGAGGCTTCAGCATGTGTCTACCGGGAGCTGGGTGCCGT	681
OY	787	GTTGCACTGCAGTGTGCTTCAGGGTGTGCTGGCGCTCAGCTGTGTGGCTGGGCGCCCACT	846
Db	682	GTTGCACTGCAGTGTGCTTCAGGGTGTGCTGGCGCTCAGCTGTGTGGCTGGGCGCCCACT	740
OY	847	GCCCTTGCAGATGTGGCCCTTCTCAGCCCTCTAGCTGTTCAGCTATGTAGCTGTGTAAAGAC	906
Db	741	GCCCTTGCAGATGTGGCCCTTCTCAGCCCTCTAGCTGTTCAGCTATGTAGCTGTGTAAAGAC	800
OY	907	TGGCTCAATGTGGGCTCAGGCGCCAGAGACCAAGACTTCTGTGGGCATAGAGCTGTGGCTGT	966
Db	801	TGGCTCAATGTGGGCTCAGGCGCCAGAGACTTCTGTGGGCATAGAGCTGTGGCTGT	859
OY	967	GGCATCTAGGGAGCGGTGTGCTGCAGTGTCTACCTACCCCTTGTGACGTGTAAAGACCAA	1026
Db	860	GGCATCTAGGGAGCGGTGTGCTGCAGTGTCTACCTACCCCTTGTGACGTGTAAAGACCAA	918
OY	1027	CGCCAGGTGCGGTGTGGGAGCGCATGTGAAGCTGTGAGAGTGAACCCCGCATGTGGACCTC	1086
Db	919	CGCCAGGTGCGGTGTGGGAGCGCATGTGAAGCTGTGAGAGTGAACCCCGCATGTGGACCTC	976
OY	1087	ACCTGG 1092	
Db	977	ACTGG 982	
RESULT	6		
LOCUS	AL529906		
DEFINITION	AL529906 LTI.NFL001_NBC4 Homo sapiens cdna clone CS0DD005YL19	EST	13-FEB-2001
ACCESSION	AL529906		
VERSION	AL529906.1	GI:12793399	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
FEATURES	Genoscope - Centre National de Sequencage		
source	BP 191 91006 EVRT cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
	Location/Qualifiers		
	1..921		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0DD005YL19"		
	/clone_1lb="LTI.NFL001_NBC4"		
	/sex="male"		

```

/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(UT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      154 a      292 c      269 g      193 t      13 others
ORIGIN

Query Match      50.1%; Score 823; DB 10; Length 921;
Best Local Similarity 98.1%; Pred. No. 2,4e-188;
Matches 831; Conservative 10; Mismatches 5; Indels 1; Gaps 1;

OY 187 CAGGACTGAAGCTTCAGATGCTGACAGAGACCTGGGGCATCAGCCCTCCAGCAA 246
    |||||||
Db 76 CAGAGCTGAAGCTTCAGATGCTGACAGAGACCTGGGGCATCAGCCCTCCAGCAA 135

OY 247 ATGTTGGCTCAGGACCGGGGCTGTGTTACTCTCTCTTCATGACACCCCTGAGCTG 306
    |||||||
Db 136 ATGTTGGCTCAGGACCGGGGCTGTGTTACTCTCTCTTCATGACACCCCTGAGCTG 195

OY 307 GTGAAGGTTGGCTTCAGAGTTCAGAGGCTCCATGGCCAGCAGCTATGCTCTCC 366
    |||||||
Db 196 GTGAAGGTTGGCTTCAGAGTTCAGAGGCTCCATGGCCAGCAGCTATGCTCTCC 255

OY 367 AGACTGTGAGACCTCTCTCTATACCAAAATGAAAGTCCCTCTGTATTTGCAATGTTGCTG 426
    |||||||
Db 256 AGACTGTGAGACCTCTCTCTATACCAAAATGAAAGTCCCTCTGTATTTGCAATGTTGCTG 315

OY 427 GAGCCTCTGTACCTGTGCTCCCAATGATGTCGCCCTGTGCTGCTTCAAGACCTTAC 486
    |||||||
Db 316 GAGCCTCTGTACCTGTGCTCCCAATGATGTCGCCCTGTGCTGCTTCAAGACCTTAC 375

OY 487 CGCTTCACTGGACCATGATGGATGCTGTGTAAGATCGTAGAGGACAGAGGACACGAGAC 546
    |||||||
Db 376 CGCTTCACTGGACCATGATGGATGCTGTGTAAGATCGTAGAGGACAGAGGACACGAGAC 435

OY 547 CTCTGAGACGGCTCCCGCCACCTGTGATGACTGTGACAGCTACCGCATCTACTTC 606
    |||||||
Db 436 CTCTGAGACGGCTCCCGCCACCTGTGATGACTGTGACAGCTACCGCATCTACTTC 495

OY 607 ACTGCTATGACCAACTGAAGGCTTCTCTGTGTGCTGACAGCCTGACCTTACCTTAC 666
    |||||||
Db 496 ACTGCTATGACCAACTGAAGGCTTCTCTGTGTGCTGACAGCCTGACCTTACCTTAC 555

OY 667 GCACCCATGATGGTGGGGGCGCTGGCCGCTTGGGACACGATGATGATGAGCCCTG 726
    |||||||
Db 556 GCACCCATGATGGTGGGGGCGCTGGCCGCTTGGGACACGATGATGATGAGCCCTG 615

OY 727 GAGCTTATGCGGAAAGCTCAGAGCTCAGCATGTGTGTCACCGGAGACTGGTGCCTGT 786
    |||||||
Db 616 GAGCTTATGCGGAAAGCTCAGAGCTCAGCATGTGTGTCACCGGAGACTGGTGCCTGT 675

OY 787 GTTGAAGTGAAGTGGCTCAGAGGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
    |||||||
Db 676 GTTGAAGTGAAGTGGCTCAGAGGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735

OY 847 GCCCTTGAGAGTGGCTTCTCAGACCTGTACTGCTCACTAATGAGCTGGTGAAGAGC 906
    |||||||
Db 736 GCCCTTGAGAGTGGCTTCTCAGACCTGTACTGCTCACTAATGAGCTGGTGAAGAGC 795

OY 907 TGGCTCAATGGGCTCAGAGCCGAGAGACGACTTCTGTGGCATTAGCTTTGTGGCTGT 966
    |||||||
Db 796 TGGCTCAATGGGCTCAGAGCCGAGAGACGACTTCTGTGGCATTAGCTTTGTGGCTGT 854

OY 967 GGCATCTCAGGAGCGGTGCTGCACTGCTACCTCTTGACAGTGTAAAGACCAA 1026

```

```

|||||
Db 855 GGCATCTCAGGAGCGGTGCTGCACTGCTACTACCTTTGACGTGTAAARACCAA 914
OY 1027 CGCCAGG 1033
Db 915 CGCCAGG 921

RESULT 7
AL516722 925 bp mRNA EST 13-FEB-2001
LOCUS AL516722 LTI_NFL011_NBC1 Homo sapiens cDNA clone CS0DA007YA10 5
DEFINITION prime, mRNA sequence.
ACCESSION AL516722
VERSION AL516722.1 GI:12780215
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 925)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..925
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DA007YA10"
/clone_id="LTI_NFL011_NBC1"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(UT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      153 a      294 c      279 g      192 t      7 others
ORIGIN

Query Match      49.5%; Score 814; DB 10; Length 925;
Best Local Similarity 98.5%; Pred. No. 3.5e-186;
Matches 847; Conservative 5; Mismatches 5; Indels 3; Gaps 3;

OY 187 CAGGACTGAAGCTTCAGATGCTGACAGAGACCTGGGGCATCAGCCCTCCAGCAA 246
    |||||||
Db 68 CAGGACTGAAGCTTCAGATGCTGACAGAGACCTGGGGCATCAGCCCTCCAGCAA 127

OY 247 ATGTTGGCTCAGGACCGGGGCTGTGTTACTCTCTCTTCATGACACCCCTGAGCTG 306
    |||||||
Db 128 ATGTTGGCTCAGGACCGGGGCTGTGTTACTCTCTCTTCATGACACCCCTGAGCTG 187

OY 307 GTGAAGGTTGGCTTCAGAGTTCAGAGGCTCCATGGCCAGCAGCTATGCTCTCC 366
    |||||||
Db 188 GTGAAGGTTGGCTTCAGAGTTCAGAGGCTCCATGGCCAGCAGCTATGCTCTCC 247

OY 367 AGACTGTGAGACCTCTCTATACCAAAATGAAAGTCCCTCTGTATTTGCAATGTTGCTG 426
    |||||||
Db 248 AGACTGTGAGACCTCTCTATACCAAAATGAAAGTCCCTCTGTATTTGCAATGTTGCTG 307

OY 427 GAGCCTCTGTACCTGTGCTCCCAATGATGTCGCCCTGTGCTGCTTCAAGACCTTAC 486

```

```

|||||
Db 308 GAGCTCTGACCTGTGCCCCAAATGATGCCGCTGTGTGACACTGTTCAAGACCTTACC 367
Qy 487 CGCTTCACTGGCACCATTGATGCTTGTGAAGATCGTAGAGCAGAGGACCAAGAGACC 546
Db 368 CGCTTCACTGGCACCATTGATGCTTGTGAAGATCGTAGAGCAGAGGACCAAGAGACC 427
Qy 547 CTCTGAGACGGGCTCCCGCCACCTGTGATGATGCTGACAGTGGACAGCCGATCTACTTC 606
Db 428 CTCTGAGACGGGCTCCCGCCACCTGTGATGATGCTGACAGTGGACAGCCGATCTACTTC 487
Qy 607 ACTGCTATGACCAACTGAAAGCCCTTCTGTGTGTGATGAGCCCTGACCTTGAACCTTAC 666
Db 488 ACTGCTATGACCAACTGAAAGCCCTTCTGTGTGTGATGAGCCCTGACCTTGAACCTTAC 547
Qy 667 GCACCCATGATGCTGGGGGCTGGCCGCTTGGGACACCTGATGATGATGATGATGATGATG 726
Db 548 GCACCCATGATGCTGGGGGCTGGCCGCTTGGGACACCTGATGATGATGATGATGATGATG 607
Qy 727 GAGCTTATGCGGACAAAGCTCAGAGCTCAGCATGTCGTACCGGAGACTGGTGCTGT 786
Db 608 GAGCTTATGCGGACAAAGCTCAGAGCTCAGCATGTCGTACCGGAGACTGGTGCTGT 667
Qy 787 GTTGCAACTGCAATGCTCAGAGGTGCTGAGGCTCCTACTGTGCTGGGCTGGGGCCCACT 846
Db 668 KTTGCAACTGCAATGCTCAGAGGTGCTGAGGCTCCTACTGTGCTGGGCTGGGGCCCACT 726
Qy 847 GCCCTTGAGATGTCCTCTTCAAGCCCTTGTACGTCGTCATGATGATGATGATGATGATGATG 906
Db 727 GCCCTTGAGATGTCCTCTTCAAGCCCTTGTACGTCGTCATGATGATGATGATGATGATGATG 786
Qy 907 TGGCTCAATGAGGCTCA-GGCCGAAGAGCAGACTTCTGTGGGCAATGAGCTTGTGTGCTG 965
Db 787 TGGCTCAATGAGGCTCAAGGCGCGAAGAMCAGACTTCTGTGGGCAATGAGCTTGTGTGCTG 846
Qy 966 TGGCATCTCAGAGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1025
Db 847 TGGCATCTCAGAGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 905
Qy 1026 AGCCAGGTGCTGCTGGGAG 1045
Db 906 AGCCAGGTGCTGCTGGGAG 925

```

```

RESULT 8
AL528266 911 bp mRNA EST 13-FEB-2001
LOCUS AL528266 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC023YA13 5
DEFINITION prime, mRNA sequence.
ACCESSION AL528266
VERSION AL528266.1 GI:12791759
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 911)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. 911
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC023YA13"
/clone_11b="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"

```

```

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a Notti-Oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 149 a 288 c 275 g 188 t 11 others
ORIGIN
Query Match 49.5%; Score 812.8; DB 10; Length 911;
Best Local Similarity 98.8%; Pred. No. 6.9e-186;
Matches 814; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
Qy 187 CAGGACTGAAAGCTTCAAGATGGCTGACAGAGACCCCTGCGGCAATCAGCCCTCCAGCAA 246
Db 88 CAGGACTGAAAGCTTCAAGATGGCTGACAGAGACCCCTGCGGCAATCAGCCCTCCAGCAA 147
Qy 247 ATGTGGCTTACGACACCGGGGCTGTGTACCTCTCTTTCATGACACCCCTGAGCTG 306
Db 148 ATGTGTGCTTACGAGMACCGGGGCTGTGTACCTCTCTTTCATGACACCCCTGAGCTG 207
Qy 307 GTGAAGGTGGCTTGCATGCTCAGCGGCGCTCCATGCGACAGAGCTGATGCTTCTTCC 366
Db 208 GTGAAGGTGGCTTGCATGCTCAGCGGCGCTCCATGCGACAGAGCTGATGCTTCTTCC 267
Qy 367 AGACTGTGAGAGCTTCTCTATACCAATGAAAGTGCCTCTGTATTCATGATGTGCTG 426
Db 268 AGACTGTGAGAGCTTCTCTATACCAATGAAAGTGCCTCTGTATTCATGATGTGCTG 327
Qy 427 GAGCTCTGCTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
Db 328 GAGCTCTGCTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 387
Qy 487 CGCTTCACTGCGACCATGATGATGCTGTGAAGATCGTAGAGCAGAGGCGCACCGAGACC 546
Db 388 CGCTTCACTGCGACCATGATGATGCTGTGAAGATCGTAGAGCAGAGGCGCACCGAGACC 447
Qy 547 CTCTGAGAGCGCTCCCGCCACCTGTGTGATGATGATGATGATGATGATGATGATGATG 606
Db 448 CTCTGAGAGCGCTCCCGCCACCTGTGTGATGATGATGATGATGATGATGATGATGATG 507
Qy 607 ACTGCTATGACCAACTGAAAGCCCTTCTGTGTGTGATGAGCCCTGACCTTGAACCTTAC 666
Db 508 ACTGCTATGACCAACTGAAAGCCCTTCTGTGTGTGATGAGCCCTTGAACCTTGAACCTTAC 567
Qy 667 GCACCCATGATGCTGGGGGCTGGCCGCTTGGGCAACCTGATGATGATGATGATGATGATG 726
Db 568 GCACCCATGATGCTGGGGGCTGGCCGCTTGGGCAACCTGATGATGATGATGATGATGATG 627
Qy 727 GAGCTTATGCGGACAAAGCTCAGAGCTCAGCATGTCGTACCGGAGACTGGTGCTGT 786
Db 628 GAGCTTATGCGGACAAAGCTCAGAGCTCAGCATGTCGTACCGGAGACTGGTGCTGT 687
Qy 787 GTTGCAACTGCAATGCTCAGAGGTGCTGAGGCTCCTACTGTGCTGGGCTGGGGCCCACT 846
Db 688 GTTGCAACTGCAATGCTCAGAGGTGCTGAGGCTCCTACTGTGCTGGGCTGGGGCCCACT 747
Qy 847 GCCCTTGAGATGTCCTTCTCAGCCCTTGTACTGCTTCAACTATGAGCTGTGAAGAGC 906
Db 748 GCCCTTGAGATGTCCTTCTCAGCCCTTGTACTGCTTCAACTATGAGCTGTGAAGAGC 807
Qy 907 TGGCTCAATGAGGCTCAGGCGGAGGACAGACTTCTGTGGGCAATGAGCTTGTGGCTGT 966
Db 808 TGGCTCAATGAGGCTCAGGCGGAGGACAGACTTCTGTGGGCAATGAGCTTGTGGCTGT 867
Qy 967 GGCATCTCAGGAGAGGTGCTGACGATGCTGACTTACCTTTGA 1010
Db 868 GGCATCTCAGGAGAGGTGCTGACGATGCTGACTTCCCTTTGA 911

```

AC007934	RESULT	9
LOCUS	AK007934	1450 bp mRNA HTC 05-JUL-2001
DEFINITION	Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:181006ZG21, full insert sequence.	
ACCESSION	AK007934	
VERSION	AK007934.1	gi:12841807
KEYWORDS	CAP trapper.	
SOURCE	Mus musculus (strain:C57BL/6J) 10 day old male pancreas cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library clone:181006ZG21.	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	1 (bases 1 to 1450)	
TITLE	Carninci, P. and Hayashizaki, Y.	
JOURNAL	High-efficiency full-length cDNA cloning	
MEDLINE	Methods in enzymology. 303, 19-44 (1999)	
PUBMED	99279253	
REFERENCE	10349636	
AUTHORS	2 (bases 1 to 1450)	
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
MEDLINE	Genome research. 10 (10), 1617-1630 (2000)	
PUBMED	20499374	
REFERENCE	11042159	
AUTHORS	3 (bases 1 to 1450)	
TITLE	Shibata, K., Itoh, M., Alizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Suni, N., Ishi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, K., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachi, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
JOURNAL	RIKEN integrated sequence analysis (RISA) system-384-format	
MEDLINE	Genome research. 10 (11), 1757-1771 (2000)	
PUBMED	20530913	
REFERENCE	11076861	
AUTHORS	4 (bases 1 to 1450)	
TITLE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	
JOURNAL	Functional annotation of a full-length mouse cDNA collection	
MEDLINE	Nature 409, 685-690 (2001)	
PUBMED	5 (bases 1 to 1450)	
REFERENCE	Adachi, J., Alizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Horii, F., Imoto, K., Ishi, Y., Itoh, M., Iwata, M., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.	
JOURNAL	Direct Submission	
TITLE	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	
COMMENT	Further details	
	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.	

Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGATCCAGAAGCTCTTTTGTTTTTTTTAA 3'], cDNA was prepared by using triethanol thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 20.0. Second strand cDNA was prepared with the primer adapter of sequence[5'-GAGAGAGAGGGCCCCCAATTATTCATTCAGTAATAAATTAATACCCCCCCC 3'}. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

```

FEATURES
    Source
        Location/Qualifiers
            1..1450
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /db_xref="MGD:MGI:1895562"
                /db_xref="MGI:MGI:1915316"
                /clone="181062G21"
                /sex="male"
                /tissue_type="pancreas"
                /clone_id="Riken full-length enriched mouse cdna library"
                /dev_stage="10 day old"
                98..1177
                    /note="putative"
                    /codon_start=1
                    /protein_id="BAB25360.1"
                    /db_xref="GI:12841808"
                    /translation="MDDDPGGISPILOOMVASGAVYTSLEMTPLDVVKYRLDSORP
                    SATSEILTPSRMSISTYKSSSAQSREKLLXNGYLEIPXLPNGRCARWQDPH
                    RFETLTLFAIVIHHEGRITLMISGLPIALVMVTVAITAFATIDOLAKFEGSLTISL
                    LYAPMVGLARMGTVIVSPLELVFRKLQOHVSRYELASVGAAYVGSGSLIMLG
                    WGPALRDVPFSALYWNVELYKSMLSGLRPKDOTSVGISFVAGISGMVAATLTLPPE
                    DVPTACROSLGAVEAVKPKPRVDSTWILLRLRIRASBSGTGLPAGFLPRIKAAPSC
                    AIMSYTEFGKSFQRINQEQPLNR"
BASE COUNT      267 a      456 c      402 g      325 t
ORIGIN
Query Match          49.1%; Score 806.6; DB 12; Length 1450;
Best Local Similarity 83.0%; Pred. No. 2.3e-184;
Matches 942; Conservative 0; Mismatches 169; Indels 24; Gaps 1;
```

Qy	153	CCCTTGAGATTGGTGGCTCTGTTTGAGGTTGGTTTGTTTAGACACTGAAGCTTCAAGATGGCTGA	212
Db	46	CCCTTGAGATTGGTGGCTCTGTTTGAGGTTGGTTTGTTTAGACACTGAAGCTTCAAGATGGCTGA	105
Qy	213	CCAGGACCCTCGGGGCATCAGGCCCTCCAGCAATATGGTGGCTTCAGGACACCGGGGGCTGT	272
Db	106	TGAGGACCCTGGGGCATTAAGCCCCCTTCAGCAATATGGTGGCTTCAGGACACCGGGGGCTGT	165
Qy	273	GGTATACCTCTCTCTTCATGACAACCCCTGGAGCTGGTGAAGTTGCGCTGCACTCACCG	332
Db	166	GGTACACCTCTCTTCATGACAACCCCTGGAGCTGGTGAAGTTGCGCTGCACTCACAG	225
Qy	333	GCCCTCCATGGCAGCAGAGCTGATGCTTCCTCCAGATGATGGAGACCTCTCCTATTACAA	392
Db	226	AACCTCGGCAACAGACGAAATGACAACCTCCAGATTCCTGGAATCTCTCTACACCAA	285
Qy	393	AT-----GGAAGTGCCTCTGATTATGCAATGGTGTCTTGGGA	428
Db	286	ATCATCTCCGCTCCAGTACGTCCTCCAGGAGAAATGCTCTCTATTACGCAATGAGATCTTGA	345
Qy	429	GCCCTCTGACCTGGGCCCAATATGGTGGCCCGCTGGGCCCACTGGTTTCAAGAACCTACCGG	488
Db	346	GCCCTCTGACCTGGGCCCAATATGGTGGCCCGCTGGGCCCACTGGTTTCAAGAACCTACCGG	405
Qy	489	CTTCACTGGACCAATGATGCTCTGTAAMAATGATGTAAGCAGACAGGACAGGACAGACCT	548
Db	406	GTTCACTGGACCACTTGATGCTCTGTAABAATGATGTAAGCAGACAGGACAGGACAGACCT	465
Qy	549	CTGAGAGGGGCTCCCGCCACCTCTGGTGTATCATCTGTGCGACACTACCGCATCTACTTAC	608
Db	466	GTTGAGAGGGGCTCCCGCCACCTCTGGTGTATCATCTGTGCGACACTACCGCATCTACTTAC	525

Accession	Source	Organism	Reference	Authors	Journal	Comment
609	TCGCATATGACCAACTGAAGGCGCTTCCTGTTGTGCTGAGCCCTGACCTTCGATCCTTACGC	668				
526	TGCTTACGACCAACTCAAGGCGCTTCCTGTTGTGCTGAGCTTCGATCCTTCCTTACGC	585				
669	ACCCATGGTGGCTGGCGGCGCTGGCCCGCTTGGGCAACCGTACTGTGATCAGCCCCCTGGA	728				
586	ACCCATGGTGGCTGGCGGCGCTGGCCCGCTTGGGCAACCGTACTGTGATCAGCCCCCTGGA	645				
729	GCTTATGCGGACAAAGCTGACAGGCTCACATGTGTCGACAGGAGCTGGGCTGGCTGTGT	788				
646	GCTGTGGGACCAAGCTGACAGGCTCACATGTGTCGACAGGAGCTGGGCTGGCTGTGT	705				
789	TCGAATCTAGTGGCTCAGGCTGGCTGGCGCTCAGTGTGGCTGGGCTGGGCCCCACATGC	848				
706	TCAACCTCCGGTGATCTCAGGGGTGGCTGGCGCTCTCTGTGGCTGGGCTGGGCTCCACAGC	765				
849	CCTTGCAGATGATGCGCTTCAGCCCTGACCTTACCTGGTTCAATATGATGCTGGTGAAGAGCTG	908				
766	TCTTGCAGATGATGCGCTTCAGCCCTTCTAGCTCTTACTGGTTCACTACGATTTGGTGAAGAGCTG	825				
909	GCTCAATGGGCTCAGGCGCGAAGGACACAGCTTCTGTGGGCACTGAGCTTTGTGGCTGGTG	968				
826	GCTGAGCGGAGCTGAGAACCCAAAGACCAACATCGGTGGGCACTGAGCTTCGTGGCGGTGG	885				
969	CATCTCAGGAGCGGGGTGGCAGTGTGACATCTACCTTACCGTGAAGACCAACAGC	1028				
886	CATCTCAGGAGTGGGTGGCAGTGTGACATCTACCTTACCGTGAAGACCAACAGC	945				
1029	CGAGTTCGCTCTGGGAGCGATGAGAGGCTGTAGAGTGAACCCCTGCATGTGAGACTGCAC	1088				
946	ACAGATGTCTACCTGGGAGCGGTGAGAGCTGTGAGATGAGAGCGGCCACAGATGAGACTGCAC	1005				
1089	CTGGTCTGCTCTGGGAGAGATCCGGGCGGAGTGGGCAACCAAGGACTCTTTGCAAGCTT	1148				
1006	CTGGTCTGCTCTGGGAGAGATCCGGGCGGAGTGGGCAACCAAGGACTCTTTGCAAGCTT	1065				
1149	CCTTCTCGGATCAATCAAGGCTGGCCCTCTCTGTCGATCATGATACACACTTATGAGTT	1208				
1066	CCTTCTCGGATCAATCAAGGCTGGCCCTCTCTGTCGATCATGATACACACTTATGAGTT	1125				
1209	CGGCAAAAGCTTCTTCAGAGGCTGAAACCAAGACCGGCTTGGGCGGCTGGAAG	1263				
1126	TGGCAAAAGCTTCTTCAGAGGCTGAAACCAAGACCGGCTTGGGCGGCTGGAAG	1180				
RESULT 10						
LOCUS	816 bp	mRNA	EST	21-MAR-2001		
DEFINITION	602533758F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:461319 5',					
ACCESSION	BC470240					
VERSION	BC470240.1					
KEYWORDS	EST.					
ORGANISM	human.					
SOURCE	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.					
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)					
	Unpublished (1999)					
	Contact: Robert Strausberg, Ph.D.					
	Email: cga@bbs-remail.nih.gov					
	Tissue Procurement: ATCC					
	CDNA Library Preparation: Ling Hong/Rubin Laboratory					
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)					
	DNA Sequencing by: NIH Intramural Sequencing Center					
	Clone distribution: MGC clone distribution information can be					
	found through the I.M.A.G.E. Consortium/LNL at:					
	http://image.lnl.gov					
	Plate: LICM460 row: c column: 16					
	High quality sequence stop: 768.					

Query Match	Best Local Similarity	Score	DB	Length	816;
Matches 811; Conservative	99.6%;	Pred. No. 1.6e-182;	Mismatches 2;	Indels 1;	Gaps 1
256	CCCTGAGCTGGTGAAGCTTCGCCCTGCACTTCAGAGGCGCCCTCATAGGCCAGGACTGA	355			
2	CCCTGAGCTGGTGAAGCTTCGCCCTGCACTTCAGAGGCGCCCTCATAGGCCAGGACTGA	61			
356	TECCTTCCTCCAGACTGAGAGCCCTCTCTATACCAAAATGGAAGTGCCTGTATTGCA	415			
62	TGCTTCCTCCAGACTGAGAGCCCTCTCTATACCAAAATGGAAGTGCCTGTATTGCA	121			
416	ATGCTGTCTGTGAGCCCTCTGTACTGTGTGCCCCAAATGTCGCCGTGTGCCACCTGGTTTC	475			
122	ATGGGTCTCTGAGGCTCTGTACTGTGTGCCCCAAATGTCGCCGTGTGCCACCTGGTTTC	181			
476	AAGACCTTACCCGGCTTCACTGGCACCATGATGATGCTGTGAAGATCGTAGGACAGAGG	535			
182	AAGACCTTACCCGGCTTCACTGGCACCATGATGATGCTGTGAAGATCGTAGGACAGAGG	241			
536	GCACGAGAACCTCTGAGAGCGGCTCCCGCCACCCCTGGTGTATGACTGTGCACGTAACG	595			
242	GCACGAGAACCTCTGAGAGCGGCTCCCGCCACCCCTGGTGTATGACTGTGCACGTAACG	301			
596	CCATCTACTTCACTGCTCATGACCAACATGAAGCTTCTGTGTGTGTGCAGCCCTACCT	655			
302	CCATCTACTTCACTGCTCATGACCAACATGAAGCTTCTGTGTGTGTGCAGCCCTACCT	361			
656	CTGACCTTACGACCAACATGATGAGTGGCGGCTGGCCGCTTGGGACACCGTATGATGA	715			
362	CTGACCTTACGACCAACATGATGAGTGGCGGCTGGCCGCTGGGACACCGTATGATGA	421			
716	TCAGCCCCCTGAGCTTATGCGGACAAAGCTGCAGGCTCAGCATGTGTGTAACCGGAGC	775			
422	TCAGCCCCCTGAGCTTATGCGGACAAAGCTGCAGGCTCAGCATGTGTGTAACCGGAGC	481			
776	TGGGTGCTGTGTTGAAGTGCAGAGGCTCAAGGTCGAGGTCGAGGTCACATGTGGCTGGCT	835			
482	TGGGTGCTGTGTTGAAGTGCAGAGGCTCAAGGTCGAGGTCGAGGTCACATGTGGCTGGCT	541			
836	GGGGCCCAACGCTCTTGAGATGTGCCCTTCTCAGCCCTGTACTGTGTTCAACTATGAGC	895			
542	GGGGCCCAACGCTCTTGAGATGTGCCCTTCTCAGCCCTGTACTGTGTTCAACTATGAGC	601			
896	TGGTGAAGAGCTGGCTCAATGGGCTTCAGGCCGCAAGACCAAGCTTGTGGGCAATGAGCT	955			
602	TGGTGAAGAGCTGGCTCAATGGGCTTCAGGCCGCAAGACCAAGCTTGTGGGCAATGAGCT	661			
956	TTGTGGCTGTGAGCAATTCAGAGGAGGTCGAGGTCAGTGTGATCTACCTTTGACGTTGG	1015			
662	TTGTGGCTGTGAGCAATTCAGAGGAGGTCGAGGTCAGTGTGATCTACCTTTGACGTTGG	721			
1016	TAAAGACCAACGCCAGGTGCTGTGGAGCATGATGAGGCTGTGAGAGT-GAACCCCTTG	1074			

```

|||||
Db 722 TAAGACCAACGCGAGGTGCTGCGAGGATGAGAGCGTGTGAGATGAAACCCCTG 781
OY 1075 CATGTGACTCCACCTGGCTGCTGCTGCGAGGA 1108
|||||
Db 782 CATGTGACTCCACCTGGCTGCTGCTGCGAGGA 815

RESULT 11
AL514846 885 bp mRNA EST 13-FEB-2001
LOCUS AL514846 LTI_NFL006.PL2 Homo sapiens cDNA clone Cl08B014ZF10 5
DEFINITION prime, mRNA sequence.
ACCESSION AL514846
VERSION AL514846.1 GI:12778339
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 885)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Source
1. 885
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site:1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dt) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : filiang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 146 a 283 c 264 g 190 t 2 others
ORIGIN
Query Match 48.5%; Score 796.4; DB 10; Length 885;
Best Local Similarity 99.4%; Pred. No. 6.2e-182;
Matches 819; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

OY 195 AAGCTTCAAGATGGCTGACGAGGACCCGCGGGGATGACGCCCTCCAGCAATGGTGGC 254
Db 64 AGCTTCAAGATGGCTGACGAGGACCCGCGGGGATGACGCCCTCCAGCAATGGTGGC 123
OY 255 CTCAGGACACGGGGGCTGTGTACCTCTCTTCATGACACCCCTGAGCTGGTGAAGT 314
Db 124 CTCAGGACACGGGGGCTGTGTACCTCTCTTCATGACACCCCTGAGCTGGTGAAGT 183
OY 315 TCGCTGCAAGTCTAGCGGGCCCTCATGGCCAGCAGCTGATGCTTCCCTCCAGACTGTG 374
Db 184 TCGCTGCAAGTCTAGCGGGCCCTCATGGCCAGCAGCTGATGCTTCCCTCCAGACTGTG 243
OY 375 GAGGCTGCTATACCAATGGAAGTGGCTCTGTATTGCAATGGTGGCTGGAGCTCT 434
Db 244 GAGGCTGCTATACCAATGGAAGTGGCTCTGTATTGCAATGGTGGCTGGAGCTCT 303
OY 435 GTACCTGTGCCAATGAGTGGCCGCTGTGCGACCTGTGTTCAAGACCCCTACCCGCTTAC 494
Db 304 GTACCTGTGCCAATGAGTGGCCGCTGTGCGACCTGTGTTCAAGACCCCTACCCGCTTAC 363

```

```

OY 495 TGGACCATGATGCTCTCTGTAAGATGTGAGGACAGGAGGACACAGACCTCTGGAG 554
Db 364 TGGACCATGATGCTCTCTGTAAGATGTGAGGACAGGAGGACACAGACCTCTGGAG 423
OY 555 CGGCTCCCGCCACACCTGTGTGATGACTGTGCCAGCTACCCCATCTACTTACCTGCTA 614
Db 424 CGGCTCCCGCCACACCTGTGTGATGACTGTGCCAGCTACCCCATCTACTTACCTGCTA 483
OY 615 TGACCACTGAAGGCTTCCTGTTGGTGTGAGGACCTGACCTGACCTGACCTGACCAT 674
Db 484 TGACCACTGAAGGCTTCCTGTTGGTGTGAGGACCTGACCTGACCTGACCTGACCAT 543
OY 675 GGTGCTGTGGCGCTGTGGCCCTTGTGGCACCCTGACTGTGATCAAGCCCTGAGACTTAT 734
Db 544 GGTGCTGTGGCGCTGTGGCCCTTGTGGCACCCTGACTGTGATCAAGCCCTGAGACTTAT 603
OY 735 GCGGCAAAAGCTGCAAGCTCAGCATGTGTGTACCGGAGAGCTGGGTGCTGTTCGAC 794
Db 604 GCGGCAAAAGCTGCAAGCTCAGCATGTGTGTACCGGAGAGCTGGGTGCTGTTCGAC 663
OY 795 TGCAGTGGCTGAGGCTGCTGGCGCTCAGTGTGGCTGGGAGCCCACTGCTGCTTCG 854
Db 664 TGCAGTGGCTGAGGCTGCTGGCGCTCAGTGTGGCTGGGAGCCCACTGCTGCTTCG 722
OY 855 AGATGTGCCCTTCTCAGCCCTGTACTGTGTTCACTATGAGCTGTGTAAGAGCTGGCTCAA 914
Db 723 AGATGTGCCCTTCTCAGCCCTGTACTGTGTTCACTATGAGCTGTGTAAGAGCTGGCTCAA 782
OY 915 TGGGTGACGGCCGAAGACAGACTTCTGTGGCATGAGCTTTGGTGGTGGCATCTC 974
Db 783 TGGGTGACGGCCGAAGACAGACTTCTGTGGCATGAGCTTTGGTGGTGGCATCTC 841
OY 975 AGGACGAGTGGCTGAGTGTGACTCTACCTTTGACGTGTAA 1018
Db 842 AGGACGAGTGGCTGAGTGTGACTCTACCTTTGACGTGTAA 885

RESULT 12
AL523298 940 bp mRNA EST 13-FEB-2001
LOCUS AL523298 LTI_NFL003.NBC3 Homo sapiens cDNA clone CS0DC001YB08 3
DEFINITION prime, mRNA sequence.
ACCESSION AL523298
VERSION AL523298.1 GI:12786791
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 940)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Source
1. 940
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CS0DC001YB08"
/clone_lib="LTI_NFL003.NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dt) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center

```



QY 1107 GATCCGGGCGGAGTGGGCGACCAAGAGACTCTTTCAGAGCTTCTCTCGGATCATCA 1166  
 DB 470 GATCCGGGCGGAGTGGGCGACCAAGAGACTCTTTCAGAGCTTCTCTCGGATCATCA 411  
 QY 1167 GGGCGCCCGCCCTGGTGCATCATGATCAGACCTATGATGATGGGCAAGAGCTTCTCA 1226  
 DB 410 GGGCGCCCGCCCTGGTGCATCATGATCAGACCTATGATGATGGGCAAGAGCTTCTCA 351  
 QY 1227 GAGGCTGACACGAGACCGGCTTCTGGGCGGCTGAAGAGGGGCAAGAGCAAGAGCCCGT 1286  
 DB 350 GAGGCTGACACGAGACCGGCTTCTGGGCGGCTGAAGAGGGGCAAGAGCAAGAGCCCGT 291  
 QY 1287 CTCCTCCACGAGTGGGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1346  
 DB 290 TTCCTCCACGAGTGGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 231  
 QY 1347 TGAGGAGAGAGGAGGCTTTCCTTCCCTCCCGGAGACAGAGCTCCAGAGGAGGCTTCC 1406  
 DB 230 TGAGGAGAGAGGAGGCTTTCCTTCCCTCCCGGAGACAGAGCTCCAGAGGAGGCTTCC 171  
 QY 1407 TCTGGGCGGCGGAT 1466  
 DB 170 TCTGGGCGGCGGAT 111  
 QY 1467 CATCATACACGAT 1526  
 DB 110 CATCATACACGAT 51  
 QY 1527 CCGT 1576  
 DB 50 CCGT 1  
 RESULT 14  
 AK007276 1371 bp mRNA HTC 05-JUL-2001  
 LOCUS Mus musculus adult male testis cDNA, RIKEN full-length enriched  
 DEFINITION library, clone:1700125C18, full insert sequence.  
 ACCESSION AK007276  
 VERSION AK007276.1 GI:12840721  
 KEYWORDS CAP trapper.  
 SOURCE Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA,  
 clone:1700125C18.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 1371)  
 AUTHORS Carninci, P. and Hayashizaki, Y.  
 JOURNAL High-efficiency full-length cDNA cloning  
 MEDLINE Methods in enzymology. 303, 19-44 (1999)  
 PUBMED 99279253  
 10349636  
 2 (bases 1 to 1371)  
 TITLE Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome research. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 11042159  
 3 (bases 1 to 1371)  
 REFERENCE Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Kono, H., Akiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M.,  
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Kashiwagi, K.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 TITLE sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome research. 10 (11), 1757-1771 (2000)

MEDLINE 20530913  
 PUBMED 11076861  
 REFERENCE 4 (bases 1 to 1371)  
 AUTHORS The RIKEN genome exploration Research Group Phase II Team and the  
 FANTOM Consortium.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409, 685-690 (2001)  
 REFERENCE 5 (bases 1 to 1371)  
 AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
 Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,  
 Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F.,  
 Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J.,  
 Kojima, Y., Kono, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T.,  
 Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y.,  
 Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H.,  
 Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
 Tanaka, T., Teijima, Y., Toya, T., Yamamura, T., Yasunishi, A.,  
 Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
 Direct Submission  
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.go.jp,  
 URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
 Fax:81-45-503-9216)  
 Please visit our web site (http://genome.gsc.riken.go.jp/) for  
 further details.  
 FEATURES  
 source  
 location/Qualifiers  
 1..1371  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /db\_xref="MGI:1896541"  
 /db\_xref="MGI:1915316"  
 /clone="1700125C18"  
 /sex="male"  
 /tissue\_type="testis"  
 /clone\_id="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="adult"  
 123..1133  
 /note="putative"  
 /codon\_start=1  
 /protein\_id="BAB24932.1"  
 /db\_xref="GI:12840722"  
 /translation="MDQDPEWGISPLQAMASGAGAVYSLFMTPLDYVVRLOSOP  
 SASELTTPSRFWSLSYKSSSALQSGKCLVCGNGLVLYCPNCTRCATFQDPT  
 RFTGTLDAFVKIVHEGTRILWGLPTVLAIVPATAIYFAVDLKAFCGSLTSD  
 LYAPVAGALARMGVTVVVSPLRFLQAOHVSYRELAISVQAAVTOGQMSLWG  
 WGPVALRDVFPFSLVWPNVELVMSLGLRPDQTSVIGSFVAGGISGAVAAATLILPF  
 DVTYTORQMSLGAIVEAVRVPKPRVDSIWLIFAESGLNLARGDSLQASSPSSRLHPV  
 PS"  
 BASE COUNT 255 a 437 c 387 g 292 t  
 ORIGIN  
 Query Match 48.0%; Score 789.4; DB 12; Length 1371;  
 Best Local Similarity 84.0%; Freq. No. 3.3e-180;  
 Matches 925; Conservative 0; Mismatches 151; Indels 25; Gaps 2;  
 QY 167 CAGGACTGAAGCTCAAGTGGCTGACAGGACCTGGGGGAGTACGCCCTCCAGCA 246  
 DB 105 CAGGATTGAATTCACAGAGTGGATATGAGACCTTGGGGCATTAGCCCTTCAGCA 164  
 QY 247 ATGTGGGCTCAGGACGAGGGGCTGTGTACCTCTCTTCATGACACCCCTGAGCGTG 306  
 DB 165 ATGTGGGCTCAGGACGAGGGGCTGTGTACCTCTCTTCATGACACCCCTGAGTGTG 224  
 QY 307 GTGAAGTTCGCTCAGTCTCAGGCGCCCTCAGTACGCGCAGGAGCTATGCTTCCTCC 366  
 DB 225 GTGAAGTTCGCTCAGTCTCAGGCGCCCTCAGTACGCGCAGGAGCTATGCTTCCTCC 284

367 AGACTGTGGAGCCCTCTCTATACCAAT-----GGAATGC 402  
 Db 285 AGATCTGGAGTCTCTCTATACCAATATCTCCGCTTACAGTCCAGGAAATGC 344  
 Oy 403 CTCCTGTATGCAATGGTCTCTGAGCCCTTGTACCTGTGCCCCAAATGGTCCCGCTGT 462  
 Db 345 CTCCTATCTCAATGGAGTCTCTGAGCCCTTGTACCTGTGCCCCAAATGGTCCCGCTGT 404  
 Oy 463 GCCACCTGGTTTCAAGACCCCTACCTGCTGACCTGACATGATGGCTTCTGTGAAGATC 522  
 Db 405 GCCACCTGGTTTCAAGACCCCTACCTGCTGACCTGACATGATGGCTTCTGTGAAGATC 464  
 Oy 523 GTGAGCACAAGAGGACACCAAGACCTTGTGAGGCGCTCCCGCAACCTGTGATGATC 582  
 Db 465 GTGAGCACAAGAGGACACCAAGACCTTGTGAGGCGCTCCCGCAACCTGTGATGATC 524  
 Oy 583 GTGACACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642  
 Db 525 GTGACACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 584  
 Oy 643 CGAGCCCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702  
 Db 585 CAGTCCCTTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 644  
 Oy 703 ACCGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 762  
 Db 645 ACCGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 704  
 Oy 763 TCGTACCGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822  
 Db 705 TCGTACCGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764  
 Oy 823 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882  
 Db 765 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 824  
 Oy 883 TTEACTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 942  
 Db 825 TTEACTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 884  
 Oy 943 GTGGGACATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1002  
 Db 885 GTGGGACATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944  
 Oy 1003 CCTTTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062  
 Db 945 CCTTTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1004  
 Oy 1063 GTGAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1122  
 Db 1005 GTGAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063  
 Oy 1123 GGCACCAAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1182  
 Db 1064 GGCACCAAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123  
 Oy 1183 GGCATCATGATCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1242  
 Db 1124 GGCATCATGATCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1183  
 Oy 1243 CGGCTTCTGGGCGGCTGAAG 1263  
 Db 1184 CAGCCTCTGGGCGGCTGAAG 1204

RESULT 15  
 AL571090 795 bp mRNA EST 16-FEB-2001  
 LOCUS AL571090.L1.NFL006.PL2 Homo sapiens cDNA clone CS0D1028YH09 3  
 DEFINITION prime, mRNA sequence.  
 ACCESSION AL571090  
 VERSION AL571090.1 GI:12928041  
 KEYWORDS EST.

SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 795)  
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
 FEATURES  
 source  
 1..795  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1028YH09"  
 /clone\_id="L1.NFL006.PL2"  
 /issue\_type="placenta"  
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-stranded cDNA was digested with Not I and  
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
 vector. Library was normalized. Library was constructed by  
 Life Technologies. Contact : Feng liang life technologies,  
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
 Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com."

BASE COUNT 167 a 241 c 231 g 142 t 14 others

Query Match 47.9%; Score 786.6; DB 10; Length 795;  
 Best local Similarity 98.1%; Pred No. 1.4e-179;  
 Matches 780; Conservative 14; Mismatches 1; Indels 0; Gaps 0;

Oy 725 TGGAGCTTATGCGGAGCAAAAGCTGACGCTGATGTCGACCGGAGCTGGTGCCT 784  
 Db 795 TGGAGCTTATGCGGAGCAAAAGCTGACGCTGATGTCGACCGGAGCTGGTGCCT 736  
 Oy 785 GTGTGGAAGTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 844  
 Db 735 GTGTGGAAGTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 844  
 Oy 845 CTGCT 904  
 Db 675 CTGCT 904  
 Oy 905 GCTGCTCAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 964  
 Db 615 GCTGCTCAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 964  
 Oy 965 GTGGGACATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1024  
 Db 555 GTGGGACATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1024  
 Oy 1025 AACGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1084  
 Db 495 AACGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1084  
 Oy 1085 CCACCTGCT 1144  
 Db 435 CCACCTGCT 1144  
 Oy 1145 GCTTCTCTCTGCT 1204  
 Db 375 GCTTCTCTCTGCT 1204  
 Oy 1205 AGTTGGCAAAAGCTTCTTCAGAGGCTGAACAGCAGCGCTTCTGGGCGGCTGAAGG 1264  
 Db 315 AGTTGGCAAAAGCTTCTTCAGAGGCTGAACAGCAGCGCTTCTGGGCGGCTGAAGG 256

```

QY 1265 GGCAGAGGCGAAGACCCGCTCTCCACGGATGGGAGAGGCGAGAGACCCAG 1324
    |||||||
Db 255 GGCAGAGGCGAAGACCCGCTCTCCACGGATGGGAGAGGCGAGAGACCCAG 196
    |||||||
QY 1325 CCAAGTGCCTTTTCTCAGCACTGAGGAGGAGGAGGCTTTTCCCTCCCGGAGCA 1384
    |:|||||
Db 195 CMAAGTGCCTTTTCTCAGCACTGAGGAGGAGGAGGCTTTTCCCTCCCGGAGCA 136
    |||||||
QY 1385 AGCTCCAGGCGAGGCTGTCCCTGAGGAGGAGGAGGCTTTTCCCTCCCGGAGCA 1444
    |||||||
Db 135 AGCTCCAGGCGAGGCTGTCCCTGAGGAGGAGGAGGCTTTTCCCTCCCGGAGCA 76
    |||||||
QY 1445 CTGCTGCTCCAGTGTGGGATCATCACTTACCCACCCCAAGTTCAAGACCAATCTT 1504
    |||||||
Db 75 CTGCTGCTCCAGTGTGGGATCATCACTTACCCACCCCAAGTTCAAGACCAATCTT 16
    |||||||
QY 1505 CCAGCTGCCCCCTTC 1519
    |||||||
Db 15 CCAGCTGCCCCCTTC 1
    |||||||

```

Search completed: April 19, 2002, 21:00:45  
 Job time: 4929 sec